



**UNIVERSIDADE FEDERAL DE PERNAMBUCO
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA
INFORMAÇÃO
MESTRADO EM CIÊNCIA DA INFORMAÇÃO**

JESSICA MONIQUE DE LIRA VIEIRA

**A CONTRIBUIÇÃO DA ORGANIZAÇÃO E DA
VISUALIZAÇÃO DA INFORMAÇÃO PARA OS SISTEMAS
DE RECUPERAÇÃO DE INFORMAÇÃO**

**RECIFE
2014**



JESSICA MONIQUE DE LIRA VIEIRA



A CONTRIBUIÇÃO DA ORGANIZAÇÃO E DA VISUALIZAÇÃO DA INFORMAÇÃO PARA OS SISTEMAS DE RECUPERAÇÃO DE INFORMAÇÃO

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Informação da Universidade Federal de Pernambuco, como requisito parcial para obtenção do título de Mestre em Ciência da Informação.

Área de Concentração: Informação, Memória e Tecnologia.

Linha de Pesquisa: Memória da Informação Científica e Tecnológica.

Orientador: Prof. Dr. Fabio Assis Pinho.

**RECIFE
2014**

Catálogo na fonte
Bibliotecária Maria Valéria Baltar de Abreu Vasconcelos, CRB4-439

V658c Vieira, Jessica Monique de Lira
 A contribuição da Organização e da Visualização da
 Informação para os Sistemas de Recuperação de Informação
 / Jessica Monique de Lira Vieira. – Recife: O Autor, 2014.
 227 p.: il.

Orientador: Fábio Assis Pinho.

Dissertação (Mestrado) – Universidade Federal de
Pernambuco. Centro de Artes e Comunicação. Ciência da
Informação, 2014.

Inclui referências.

1. Ciência da Informação. 2. Organização da Informação.
3. Recuperação da informação. 4. Sistemas de recuperação
da informação. I. Pinho, Fábio Assis (Orientador). II. Título.

020 CDD (22.ed.) UFPE (CAC 2014-1)

Dissertação de Mestrado apresentada no dia 24 de fevereiro de 2014, por JÉSSICA MONIQUE DE LIRA VIEIRA ao Programa de Pós-graduação em Ciência da Informação do Centro de Artes e Comunicação da Universidade Federal de Pernambuco, sob o título **“A CONTRIBUIÇÃO DA ORGANIZAÇÃO E DA VISUALIZAÇÃO DA INFORMAÇÃO PARA OS SISTEMAS DE RECUPERAÇÃO DE INFORMAÇÃO”** orientada pelo Prof. Dr. Fabio Assis Pinho e **aprovada** pela Banca Examinadora composta pelos professores:

Prof. Fabio Assis Pinho
Departamento de Ciência da Informação-PPGCI/UFPE

Prof. Fábio Mascarenhas e Silva
Departamento de Ciência da Informação-PPGCI/UFPE

Prof. Dr. Fabiano Ferreira de Castro
Núcleo de Ciência da Informação/UFS

Autor:

Jéssica Monique de Lira Vieira

LISTA DE ILUSTRAÇÕES

Figura 1 – Componentes de um sistema de recuperação de informação	40
Figura 2 – OC/RC e OI/RI	65
Figura 3 – Aspectos inerentes ao TDI	77
Figura 4 – Aspectos inerentes ao TTI	93
Figura 5 – Visualizações da cotação do Dólar para compra.....	116
Figura 6 – Desenho de Grafos	120
Figura 7 – Desenho de Grafos	121
Figura 8 – Desenho de árvores das áreas de conhecimento do CNPq	123
Figura 9 – Browser Hiperbólico do <i>software</i> Inxight	125
Figura 10 – <i>Fish-eye</i> menu	127
Figura 11 – Aplicação da técnica <i>perspective wall</i> desenvolvida pelo Inxight.....	128
Figura 12 – <i>Table Lens</i>	129
Figura 13 – Mapa Conceitual sobre Mapas Conceituais	132
Figura 14 – Cartografia temática para o PPG em Direito, baseada em classificação hierárquica de descritores - DENDROGRAMA	135
Figura 15 – Cartografia temática para o PPG em Direito, baseada em classificação em espaço reduzido não hierárquica de descritores – ESCALA MULTIDIMENSIONAL (MDS)	136
Figura 16 – Nuvem de <i>Tags</i>	139
Figura 17 – Desenho de grafos da Biblioteca da Universidade de Edimburgo.....	151

Figura 18 – Desenho de árvores do YIPPY.....	158
Figura 19 - Resultado de uma busca no desenho de árvores do YIPPY	160
Figura 20 - Árvore hiperbólica da Infoteca-e da Embrapa ..	166
Figura 21 - Busca realizada na árvore hiperbólica da Infoteca-e da Embrapa.....	168
Figura 22 - Descrição do conteúdo do vértice	170
Figura 23 - Mapa Conceitual (IHMC CmapTools).....	174
Figura 24 - Ícones para acesso a documentos no mapa conceitual.....	176
Figura 25 - Ícone para visualização de mapa conceitual....	177
Figura 26 - Mapa conceitual do CMEX Mars.....	178
Figura 27 - Mapa gerado pelo WEBSOM, baseado na técnica cartografia temática	184
Figura 28 - Zoom sobre uma área específica do mapa	185
Figura 29 - Nuvens de Tags do BRAPCI.....	191
Figura 30 - Nuvem de tags gerada a partir de busca por meio do termo organização da Informação	192
Figura 31 - Análise do foco das áreas de OI e VI	204

LISTA DE QUADROS

Quadro 1- Princípios da indexação	46
Quadro 2 - Relação entre OI e VI na aplicação de interface visual	200

LISTA DE ABREVIATURAS E SIGLAS

AACR	<i>Anglo-American Cataloguing Rules</i>
AACR2	<i>Anglo-American Cataloguing Rules, second edition</i>
ALA	<i>American Library Association</i>
BRAPCI	Base de Dados Referenciais de Artigos de Periódicos em Ciência da Informação
CAPES	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
CC	<i>Colon Classification</i>
CDD	Classificação Decimal de Dewey
CDU	Classificação Decimal Universal
CI	Ciência da Informação
CMEX MARS	Centro de Exploração de Marte da Nasa
CRG	<i>Classification Research Group</i>
EMBRAPA	Empresa Brasileira de Pesquisa Agropecuária
FRAD	<i>Functional Requirements for Authority Data</i>
FRBR	<i>Functional Requirements for Bibliographic Records</i>
HTML	<i>Hypertext Markup Language</i>
IFLA	<i>International Federation of Library Association and Institutions</i>
IHMC	Institute for Human & Machine Cognition
INFOTECA-E	Serviço Informação Tecnológica em Agricultura

ISBD	<i>International Standard Bibliographic Description</i>
ISKO	<i>International Society for Knowledge Organization</i>
MDS	Escala Multidimensional
OAI-PMH	<i>Open Archives Initiative - Protocol for Metadata Harvesting</i>
OC	Organização do conhecimento
OI	Organização da Informação
PD&I	Pesquisa, Desenvolvimento e Inovação
RDA	<i>Resource Description and Access</i>
RIEC	Reunião Internacional de Especialistas em Catalogação
SCIELO	<i>Scientific Electronic Library Online</i>
SRI	Sistema de Recuperação de Informação
SRIs	Sistemas de Recuperação de Informação
TDI	Tratamento Descritivo da Informação
TTI	Tratamento Temático da Informação
UNESCO	<i>United Nations Educational, Scientific And Cultural Organization</i>
VI	Visualização da Informação

Dedico aos meus familiares e a Deus!

“Sim, coisas grandiosas fez o Senhor por nós, por isso estamos alegres”.

(Salmo 126:3)

AGRADECIMENTOS

Acima de tudo, agradeço a Deus que me proporcionou a realização desta pesquisa, conduzindo em cada passo e concedendo sabedoria e forças para superar cada obstáculo até a reta final. Graças a Ele posso concluir mais uma etapa da minha vida.

Agradeço ao Professor e orientador Fabio Assis Pinho pelos ensinamentos, pela confiança, apoio e orientação no desenvolvimento da dissertação.

Aos meus familiares por todo apoio e força.

À minha mãe pelo amor, paciência e esforço empenhado para que eu pudesse me dedicar ao mestrado.

À minha irmã, que sempre esteve ao meu lado me animando e apoiando em cada etapa.

Aos meus amigos pela força e encorajamento.

Ao Programa de Pós-Graduação em Ciência da Informação da Universidade Federal de Pernambuco, pelo incentivo.

A CAPES pelo auxílio financeiro que viabilizou o desenvolvimento da pesquisa.

“A informação nos instrui sobre a natureza do mundo em que vivemos; sua história, seu futuro, seu funcionamento, nosso lugar nele, nossas ações possíveis e suas conseqüências potenciais. Nesse sentido a informação é extremamente valiosa. Ela nos oferece um valor de sobrevivência e instruções não somente para viver, mas para viver melhor”.

Brenda Dervin

RESUMO

Buscando propor estratégias que potencializem o entendimento da memória armazenada em Sistemas de Recuperação de Informação (SRIs), considera-se, neste trabalho, como alternativa, a adoção de procedimentos de Organização da Informação (OI) e Visualização da Informação (VI). A incipiência de estudos que identifiquem a OI e a VI como áreas que convergem entre si e que, interligadas, atuando em parceria, são capazes de contribuir para o aumento na precisão dos SRIs, é focada nesse trabalho enquanto problema de pesquisa. Nesse contexto, tornou-se necessário analisar e identificar a relação entre os aspectos da OI e da VI do ponto de vista da contribuição das duas áreas para os SRIs, de maneira a potencializar o entendimento da memória por eles armazenada, razão pela qual a pesquisa é caracterizada como exploratória e bibliográfica. O percurso metodológico utilizado para realização da pesquisa consiste em levantamento bibliográfico, para obtenção do embasamento científico para melhor compreensão e análise de cada área; levantamento de sistemas que fazem uso de interface de VI em sua estrutura, bem como da análise dos elementos de OI necessários para organização e representação de conteúdos a serem transmitidos em estruturas visuais. A partir disso, evidenciou-se a interligação entre as áreas abordadas, quando aplicadas em sistemas que lidam com a recuperação da informação. Desse modo, concluiu-se que a OI e a VI podem alcançar promissores resultados se atuarem em parceria para otimização na recuperação da memória armazenada em SRIs.

Palavras-chave: Organização da Informação. Visualização da Informação. Sistemas de Recuperação de Informação.

ABSTRACT

Trying to propose strategies that enhance the understanding of memory stored in Information Retrieval Systems (IRSs), this thesis considers, as an alternative, the adoption of procedures of Information Organization (IO) and Information Visualization (IV). The paucity of studies that identify the IO and VI as areas that converge with each other and that, interconnected, working together, are able to contribute to increase the accuracy of SRIs is focused in this work as a research problem. Thus, it became necessary to analyze and identify the relationship between the resources from the IO and IV from the point of view of the contribution of the two areas for SRIs, in order to enhance the understanding of memory stored by them, for that reason the research is characterized as exploratory and bibliographical. The methodological procedure used to conduct the research consists of bibliographic survey, for obtaining scientific basis for better understanding and analysis of each area; a survey of systems that make use of IV interface in its structure, as well as the analysis of the IO criteria necessary to organization and representation of content to be transmitted in visual structures. From that, it was evident the interconnection between the areas addressed when applied in systems that deal with information retrieval. Thus, it is concluded that the IO and IV can achieve promising results if they work in partnership to optimize the recovery of memory stored in IRSs.

Keywords: Information organization. Information Visualization. Information Retrieval Systems.

SUMÁRIO

1	INTRODUÇÃO	17
1.1	Justificativa	25
1.2	Objetivos.....	27
2	RECUPERAÇÃO DA INFORMAÇÃO	31
2.1	Dificuldade enfrentada pelos Sistemas de Recuperação de Informação	43
2.2	Modelos de Recuperação de Informação utilizados em SRIs	48
2.2.1	<i>Modelos Quantitativos.....</i>	<i>49</i>
2.2.2	<i>Modelos Dinâmicos.....</i>	<i>56</i>
3	ORGANIZAÇÃO DA INFORMAÇÃO	63
3.1	Tratamento da Informação	66
3.1.1	<i>Tratamento Descritivo da Informação.....</i>	<i>67</i>
3.1.1.1	Fundamentos teóricos.....	71
3.1.1.2	Processos, produtos, instrumentos	77
3.1.2	<i>Tratamento Temático da Informação.....</i>	<i>81</i>
3.1.2.1	Fundamentos teóricos.....	82
3.1.2.2	Correntes teóricas.....	88
3.1.2.3	Processos, Produtos, Instrumentos.....	92
3.2	Organização da Informação em ambiente	

Digital.....	100
--------------	-----

4 VISUALIZAÇÃO DA INFORMAÇÃO	108
------------------------------------	-----

4.1 Benefícios trazidos pelo uso da Visualização da Informação	112
4.2 Técnicas de Visualização da Informação	118
4.3 Modelos de interface de Visualização da Informação	130

5 MÉTODO	141
----------------	-----

5.1 Tipo de Pesquisa	141
5.2 Etapas da Pesquisa	142
5.2.1 Levantamento Bibliográfico.....	144
5.2.2 Descrição de técnicas e modelos de interface de VI.....	144
5.2.3 Análise de sites de SRIs	145

6 SISTEMAS QUE FAZEM USO DA VISUALIZAÇÃO DA INFORMAÇÃO	148
---	-----

6.1 AquaBrowser (Biblioteca da Universidade de Edimburgo).....	149
6.1.1 Vantagens oferecidas pela interface visual da biblioteca da universidade de Edimburgo	153
6.1.2 Análise da interface visual quanto	

<i>à expressividade e efetividade.....</i>	154
<i>6.1.3 Análise dos elementos de OI necessários para o emprego de interface visual baseada no desenho de grafos</i>	155
6.2 Mecanismo de busca YIPPY.....	156
<i>6.2.1 Vantagens oferecidas pela interface visual do mecanismo de busca YIPPY.....</i>	161
<i>6.2.2 Análise da interface visual quanto à expressividade e efetividade.....</i>	162
<i>6.2.3 Análise dos elementos de OI necessários para o emprego de interface visual baseada no desenho de árvores</i>	162
6.3 Infoteca-e da Embrapa.....	164
<i>6.3.1 Vantagens oferecidas pela interface visual da Infoteca-e da Embrapa.....</i>	169
<i>6.3.2 Análise da interface visual quanto à expressividade e efetividade.....</i>	171
<i>6.3.3 Análise dos elementos de OI necessários para o emprego de interface visual baseada no browser hiperbólico.....</i>	172
6.4 Web site do IHMC CmapTools.....	173
<i>6.4.1 Vantagens oferecidas pela interface visual do IHMC CmapTools</i>	179
<i>6.4.2 Análise da interface visual quanto à expressividade e efetividade.....</i>	180
<i>6.4.3 Análise dos elementos de OI necessários para o emprego de interface visual baseada no mapa conceitual.....</i>	181
6.5 WEBSOM - método para construção de mapas visuais	182

6.5.1 Vantagens oferecidas pela interface visual do WEBSOM	186
6.5.2 Análise da interface visual quanto à expressividade e efetividade	187
6.5.3 Análise dos elementos de OI necessários para o emprego de interface visual baseada em cartografia temática.....	188
6.6 Sistema de busca da BRAPCI	189
6.6.1 Vantagens oferecidas pela interface visual da BRAPCI	193
6.6.2 Análise da interface visual quanto à expressividade e efetividade	195
6.6.3 Análise dos elementos de OI necessários para o emprego de interface visual baseada na nuvem de tags	196
6.7 Relação entre OI e VI nos sites de SRIs que fazem uso de interface visual	198
7 CONSIDERAÇÕES FINAIS	203
REFERENCIAS	208

1 INTRODUÇÃO

A crescente produção de informações, juntamente com os avanços tecnológicos, desempenha um papel fundamental na consolidação da sociedade atual, enquanto “sociedade da informação”. Sociedade em que a informação é vista, como afirma Le Coadic (2004), como uma necessidade fundamental, ou seja, como uma necessidade vital para o desenvolvimento de, praticamente, todas as atividades humanas. É nesse cenário, no qual a informação, dotada de tamanho valor, enfrenta um crescimento na sua produção e disponibilização, que a busca por torná-la acessível se faz importante.

No entanto, face à grande quantidade de informações produzidas e dispostas, surgem também as barreiras quanto à recuperação desse conteúdo, pois se agravam os problemas de acesso e organização dos registros de informação gerados e acumulados pela sociedade (BRAZ, 2013). Sobre isso, Feitosa (2006) comenta que o crescimento da informação produzida compara-se a uma montanha que se torna cada vez mais alta e volumosa, razão pela qual se faz necessária a busca de métodos adequados para, significativamente, gerenciar a quantidade de informação disponível de forma a possibilitar seu acesso e uso.

Diante desse contexto e da necessidade de proporcionar o acesso para posterior uso da informação, torna-se salutar propor soluções para os problemas relacionados à informação e à sua recuperação. Nesse âmbito, uma das áreas preocupadas e envolvidas com essa

questão é a Ciência da Informação (CI) que, juntamente com outras áreas, é uma participante ativa e deliberada na evolução da sociedade da informação (SARACEVIC, 1995); e, como aponta Dias (2001), tem como questão básica o acesso à informação, ou a facilitação desse acesso, o que constitui uma das razões gerais de sua existência.

Assim, a CI, nas palavras de Alves et al. (2007, p. 29, grifo nosso), consiste em uma área que:

busca investigar, entre outros assuntos, a geração, a **organização**, o processamento, a **recuperação** e o uso da informação, em diversos ambientes. É, por natureza, uma área interdisciplinar [...] Além disso, está intimamente ligada com as Tecnologias de Informação e Comunicação vigentes.

Nessa perspectiva, entre as áreas estudadas pela CI, a Recuperação da Informação (RI) é vista, como menciona Saracevic (1995), como uma das áreas em que mais ocorrem as relações interdisciplinares. A Recuperação da Informação preocupa-se, portanto, em disponibilizar a informação para satisfação de uma necessidade informacional. Diante disso, torna-se oportuno evidenciar o papel dos sistemas que atuam no intuito de tornar possível, a partir de um universo de documentos disponíveis, a recuperação daqueles que venham a ser úteis para quem deles necessita, os chamados Sistemas de Recuperação de Informação (SRIs), cujo objetivo, de forma genérica, é facilitar o acesso e possibilitar a recuperação e o entendimento da memória produzida e por eles armazenada. Tais sistemas, como destaca Araújo (1995), constituem a “memória humana registrada”.

Dada a responsabilidade a eles atribuída, os SRIs buscam exercer sua função de forma eficaz, mediante realização das atividades que envolvem representação, armazenamento, e recuperação da informação, contidas em documentos, que venham a ser relevantes para uma questão formulada por um usuário (INGWERSEN, 2002). Assim, baseando-se na existência das atividades que compõem um SRI: representação, armazenamento e recuperação, pode-se inferir que há uma interação entre elas, que visa alcançar o objetivo proposto pelo sistema, ou seja, a partir de uma representação adequada e do armazenamento das informações que foram representadas, os sistemas poderão recuperar conteúdos relevantes, que satisfarão a uma necessidade informacional. Esses conteúdos, bem como suas representações, constituem-se da memória humana registrada, da qual se refere Araújo (1995).

Além disso, para que os SRIs atuem de maneira satisfatória, é fundamental que haja uma atenção maior no que tange à organização das informações armazenadas por esses sistemas, pois, como destaca Brascher e Carlan (2010), em um sistema de informação, a qualidade obtida na recuperação da informação depende, substancialmente, dos procedimentos e instrumentos utilizados para organização da informação. Dessa maneira, os padrões de organização devem ser definidos desde a concepção do sistema para permitir que a informação seja encontrada posteriormente, pois, do contrário, a recuperação do conteúdo relevante à uma necessidade pode ser comprometida, limitando, assim, o desempenho do SRI.

Nesse âmbito, tem-se a Organização da Informação (OI), que exerce uma atividade nuclear enquanto mediadora, sendo responsável por estabelecer uma ponte entre a representação e o uso da informação, caracterizando, assim, seu papel, quando se lida com objetos informacionais. Nesse sentido, a OI tem como função, resumidamente, organizar registros de informação para arranjá-los sistematicamente, de forma a possibilitar sua recuperação e posterior uso.

Para desenvolvimento de suas práticas, a OI fundamenta-se em um aspecto essencial, o tratamento da informação. O tratamento da informação, de acordo com Pinto e Molina (1993 apud EVEDOVE, 2011), consiste no processo de gerar representações documentais, as quais, inseridas no contexto de sistemas de informação, atuam como instrumentos de busca, cujo propósito corresponde à facilitação da consulta e disseminação do documento primário. Assim, tais representações funcionam como sínteses que facilitam a identificação dos documentos e se eles são relevantes a uma necessidade.

Para isso, o tratamento da informação é composto de duas vertentes: o tratamento físico e o tratamento temático. O tratamento físico compreende a descrição física dos documentos, de acordo com suas características extrínsecas, já o tratamento temático compreende a descrição de conteúdo dos documentos a partir das suas características intrínsecas. Assim, tanto o tratamento físico quanto o temático propiciam o acesso aos documentos e, conseqüentemente, às informações neles contidas, por meio de suas representações condensadas.

Encaminhando a discussão para os sistemas que lidam com informações, especificamente, em ambiente digital, torna-se importante destacar a necessidade de organização da informação que, nesse contexto, se faz presente. Já que, apesar dos significativos desenvolvimentos na produção, armazenamento e disseminação da informação, acarretados pela era digital, notadamente pelos avanços das tecnologias de informação, essa nova realidade trouxe, também, grandes desafios, principalmente no que se refere ao acesso e recuperação da informação.

Dessa forma, nesse ambiente digital, em se tratando dos procedimentos relativos ao tratamento da informação que, conforme visto anteriormente, são responsáveis pela representação documental, é importante destacar que a OI atua por meio de metadados. Os metadados são empregados para descrever e localizar documentos eletrônicos. Assim, é a partir deles que se realiza o tratamento tanto descritivo quanto temático dos itens de informação nesse ambiente, tornando possível a descrição das características extrínsecas e intrínsecas dos documentos.

Considerando essa realidade, ressalta-se o importante valor que a OI representa para os sistemas de informação, tendo em vista que a adequada organização tem o poder de interferir diretamente na recuperação de informações precisas e que, em contrapartida, a informação sem uma organização prévia pode permanecer oculta e sem uso. Reafirmando essa ideia, Svenonius (2000) enfatiza, ainda, que a eficácia de um sistema que lida com a recuperação da informação está

diretamente relacionada à inteligência adotada para sua organização.

Outro ponto a ser tratado quando se lida com a otimização na recuperação e uso da memória armazenada nos sistemas de informação refere-se a Visualização da Informação (VI). A VI pode ser vista como uma área que tem como finalidade potencializar a apropriação de informação pelo usuário, por meio de recursos gráficos conhecidos como estruturas de visualização da informação (DIAS; CARVALHO, 2007).

Nesse sentido, a VI apóia-se na representação visual como maneira de auxiliar a compreensão de informações armazenadas e que foram, previamente, organizadas, objetivando facilitar a localização e recuperação eficaz de itens que atendam a uma necessidade. Para isso, a VI faz uso de interfaces gráficas que possibilitam a interação do usuário com os sistemas, com a finalidade de maximizar o uso da informação. Dessa forma, a VI atua oferecendo um ambiente de pesquisa agradável, capaz de orientar os usuários em uma busca para as informações de seu interesse, por meio de representação visual.

No entanto, para que a VI possa exercer sua função de forma satisfatória, é preciso que os procedimentos de OI ofereçam um suporte que torne possível a adoção de representações visuais pelos sistemas (ZHANG, 2008), pois a OI se configura como área fundamental na organização e tratamento da informação, e que afeta diretamente sua recuperação; por outro lado, a VI se configura como uma poderosa aliada, capaz de potencializar o acesso, localização

e recuperação da informação organizada. Assim, tem-se a OI como um processo interno aos sistemas de recuperação de informação, porém, muitas vezes, não transparente; e a VI que, atuando mediante representação visual, visa à facilitação na apropriação das informações previamente estruturadas e organizadas.

É nesse cenário que se estabelece o tema deste trabalho, ou seja, a abordagem da interação existente entre a Organização e a Visualização da Informação e a sua influência em sistemas de recuperação de informação.

Considerando o tema abordado, é oportuno destacar que o problema que motivou a realização da pesquisa deu-se em virtude da incipiência de estudos que identifiquem a OI e a VI como áreas que convergem entre si e que, interligadas, atuando em parceria, podem potencializar os sistemas de recuperação de informação.

Uma vez que, na literatura brasileira, são incipientes os trabalhos que abordam, especificamente, que a padronização ou organização da informação pode garantir visualizações mais expressivas e, conseqüentemente, promover uma recuperação mais efetiva da memória armazenada nos SRIs, questiona-se o porquê da incipiência de trabalhos que relacionem a OI e a VI, pois se entende como áreas que possuem aproximação e que, se estudadas em conjunto, acarretariam promissores resultados no contexto da recuperação da informação.

É exatamente nesse contexto que se estabelece o propósito da pesquisa, isto é, analisar, a partir de revisão de literatura, como a OI, enquanto área que se preocupa com a

organização e tratamento da informação, pode subsidiar a construção de interfaces de VI, cujo objetivo é representar informações de forma gráfica, facilitando o acesso e recuperação de conteúdos informacionais.

Sob essa ótica, entende-se como necessária a busca por uma maior familiaridade com as temáticas aqui estudadas. Para isso, buscou-se, mediante a exploração de referencial teórico, fornecer o embasamento científico necessário, a fim de melhor compreender e analisar os aspectos que envolvem a aplicação da Organização da Informação e da Visualização da Informação em SRIs, razão pela qual a pesquisa se caracteriza como exploratória e bibliográfica.

Nesse âmbito, realizou-se um levantamento bibliográfico nos principais periódicos e eventos da área de CI, com exceção de trabalhos advindos da área da Tecnologia e da Ciência da Computação, que foram necessários para o desenvolvimento da pesquisa; e em bases de dados, destacando especificamente, a BRAPCI e CAPES, e o portal de periódicos SCIELO. A partir do levantamento bibliográfico, tornou-se possível um aprofundamento nos aspectos relevantes de cada área.

Em sequência, foi realizado um levantamento de SRIs na Web que adotam visualizações na interface com o usuário, com o intuito de identificar como a OI pode atuar na organização e representação de conteúdos a serem transmitidos em estruturas visuais, permitindo, dessa forma, uma visualização prática acerca da aplicação de OI e VI em SRIs.

1.1 Justificativa

Diante da tentativa de propor estratégias que potencializem o entendimento da memória armazenada em sistemas de recuperação de informação, a justificativa aqui pleiteada baseia-se na viabilidade da adoção de procedimentos de Organização da Informação e da aplicação de interfaces de Visualização da Informação, como alternativas capazes de contribuir para o aumento na precisão desses sistemas.

Assim, a importância da pesquisa decorre do fato de que os sistemas detentores de informação e da memória humana registrada, como citou Araújo (1995) anteriormente, têm a missão de tornar acessível o conteúdo neles armazenado, de maneira que tal conteúdo possa ser amplamente utilizado. No entanto, esses sistemas são afetados pelo fato de haver dificuldades na recuperação da informação que comprometem seu desempenho. É nesse contexto, então, que entra em cena a OI e a VI como áreas que assumem uma relação direta com a recuperação da informação.

Essa afirmação pode ser justificada, sob o ponto de vista da OI, pela declaração de Naves e Kuramoto (2006) ao evidenciar que se a informação for disposta de maneira organizada, será mais provável, em tese, localizar e recuperar o que se procura. A OI atua no tratamento da informação, permitindo sua identificação e representação, o que interfere, conseqüentemente, em sua recuperação. Dessa maneira, como menciona Medeiros (2010), a

informação é passível de tratamento e organização adequados para que não interfira negativamente em sua recuperação. Diante disso, entende-se como relevante que a OI esteja associada aos sistemas que objetivam a recuperação de informações.

Sob o ponto de vista da VI, é oportuno destacar que a atuação dessa área está direcionada ao acesso, localização e recuperação da informação, ou seja, atuando na última etapa que compreende um SRI, a etapa da recuperação da informação. Nesse âmbito, a VI propõe a utilização de recursos visuais no intuito de facilitar a compreensão acerca das informações dispostas, principalmente, em meios digitais, tendo em vista que os usuários, ao se submeterem a um sistema, por diversas vezes, apresentam um grau de incerteza quanto ao que informar no momento da consulta, bem como na seleção do que é pertinente para atender a sua necessidade, diante das grandes listas de documentos retornados.

Nesse contexto, a VI, por meio de interfaces visuais, oferece a possibilidade de navegação por entre as informações dispostas como uma maneira mais amigável de se visualizar e localizar itens úteis, quando comparado com a consulta tradicional oferecida por boa parte dos sistemas, haja vista que interfaces visuais efetivas permitem observar, manipular, pesquisar, navegar, explorar, filtrar, descobrir, entender e interagir com grandes volumes de informações mais rapidamente e mais efetivamente para descobrir padrões ocultos. (ESTIVALET, 2000). Com isso, tem-se o importante papel que a VI pode exercer no intuito de

contribuir para a recuperação de informações em SRIs.

Nessa perspectiva, a análise dessa pesquisa, em âmbito geral, torna-se oportuna ao propor a associação entre a OI e a VI, pois se entende que, para que as interfaces visuais auxiliem na recuperação de informações de maneira eficaz, se faz necessária a organização prévia dessas informações.

Em síntese, quando um sistema possui suas informações previamente organizadas e estruturadas, a utilização de uma interface visual pode potencializar a interação do usuário com o sistema, auxiliando na busca e obtenção das informações de seu interesse e, também, na dedução de novos conhecimentos a partir do que está sendo apresentado.

1.2 Objetivos

O objetivo geral e os objetivos específicos que delineiam esta pesquisa são descritos a seguir:

Objetivo Geral

Analisar a relação entre os aspectos da Organização da Informação e da Visualização da Informação do ponto de vista da contribuição das duas áreas para os Sistemas de Recuperação de Informação, de maneira a potencializar o entendimento da memória por eles armazenada.

Objetivos Específicos

- Identificar os aspectos oriundos das áreas de Organização da Informação, Visualização da Informação e Sistemas de Recuperação de Informação;
- Identificar e analisar relações entre a Organização da Informação e a Visualização da Informação e sua aplicabilidade em Sistemas de Recuperação de Informação;
- Levantar SRIs que fazem uso de Visualização da Informação;
- Analisar os SRIs levantados, quanto aos elementos de organização da informação que são necessários para o emprego de interfaces gráficas;
- Discutir, com base na análise realizada, acerca das implicações da organização da Informação na Visualização da Informação.

Dessa forma, esta dissertação está organizada em sete seções. Assim, após a apresentação dos elementos introdutórios da pesquisa, que incluem o tema, a problemática, a justificativa e os objetivos, tem-se, em seguida, a seção 2 - RECUPERAÇÃO DA INFORMAÇÃO, que contextualiza a recuperação da informação e os sistemas

que trabalham para tornar essa recuperação possível, isto é, os SRIs, destacando seus objetivos, atividades e dificuldades por eles enfrentada. Destacam-se, ainda, os modelos de recuperação da informação (quantitativos e dinâmicos) através dos quais os SRIs realizam a localização e recuperação da informação que atenda ao usuário.

Levando em conta que, para recuperar informações relevantes é preciso uma organização prévia dessas informações, na seção 3 – ORGANIZAÇÃO DA INFORMAÇÃO, abordam-se os conceitos e práticas da OI, focando o tratamento da informação e seus dois processos: tratamento descritivo e tratamento temático, seus fundamentos teóricos, bem como os processos, produtos e instrumentos que compõem a sua essência. Nessa seção abordam-se, também, os preceitos da organização da informação em ambiente digital. Dando continuidade, a seção 4 – VISUALIZAÇÃO DA INFORMAÇÃO trata da Visualização da Informação, destacando seu propósito, os benefícios trazidos pelo uso de visualizações, bem como o levantamento de algumas técnicas e modelos de interface de VI utilizadas, com o intuito de representar graficamente informações em SRIs.

A seção seguinte, seção 5 – MÉTODO, apresenta os procedimentos metodológicos utilizados para realização da pesquisa, ou seja, a metodologia empregada para o levantamento bibliográfico, descrição e análise de técnicas e modelos de interface e análise de sites de SRIs, sendo traçados, assim, os caminhos percorridos para construção de cada etapa.

Em seguida, na seção 6 – SISTEMAS QUE FAZEM USO DA VIUALIZAÇÃO DA INFORMAÇÃO, tem-se um levantamento de sites na Web que fazem uso de técnicas e modelos de interface de VI, permitindo uma visão prática da utilização de interfaces gráficas em SRIs. Nessa seção apresenta-se, também, uma análise de como deveriam estar organizadas e estruturadas as informações nos sistemas e quais instrumentos de OI seriam necessários para tornar possível a aplicação eficaz de cada técnica e modelo de VI, evidenciando, dessa forma, a relação entre OI e VI nos sites de SRIs. E, por fim, na seção 7 – CONSIDERAÇÕES FINAIS, têm-se as considerações finais acerca da pesquisa realizada e das contribuições identificadas, retomando o objetivo inicialmente proposto; seguindo-se das referências.

2 RECUPERAÇÃO DA INFORMAÇÃO

A necessidade de informação é vista, segundo Le Coadic (2004), como uma necessidade física fundamental, ou seja, necessita-se de informação para construir, solucionar, preencher lacunas, esclarecer, informar e formar o ser humano desde as meras atividades cotidianas. Sendo assim, é nesse contexto que se insere a RI, como área que se preocupa com o uso da informação para satisfação de uma necessidade informacional, visando tornar acessível a informação para quem dela necessita. Diante de tal responsabilidade, a RI pode ser considerada uma das áreas essenciais no campo da Ciência da Informação. Como explicitou Saracevic (1996), a Recuperação da Informação é o maior ramo da CI.

Nesse momento, é importante esclarecer que a informação refere-se à registrada em um suporte, passível de ser transmitida e socializada. Como aborda Lê Coadic (2004, p.4):

A informação é um conhecimento inscrito (registrado) em forma escrita (impressa ou digital), oral ou audiovisual, em um suporte. A informação comporta um elemento de sentido. É um significado transmitido a um ser consciente por meio de uma mensagem inscrita em um suporte espacial-temporal [...].

Assim, a recuperação da informação é realizada a partir de uma informação registrada, que carrega um sentido, e que é capaz de ser socializada.

Adequando a recuperação da informação à realidade atual, fortemente influenciada pelo intensivo uso de tecnologias de comunicação e informação, que tem transformado o novo século em um mundo cada vez mais ligado em rede de informação global, em que milhões de pessoas se conectam produzindo e recuperando informação em tempo real, faz surgir novas formas de pensar, interagir e viver. Esse desenvolvimento trouxe como consequências da explosão da informação, caracterizada pelo aumento da quantidade de informações produzidas e armazenadas, a diversidade e possibilidade de reprodução sem limites (FURTADO, 2010).

Neste universo intensivo de tecnologias de informação e comunicação, e em meio a essa gama de informações, os sistemas eletrônicos, que hoje permeiam e mesmo viabilizam virtualmente quase todas as atividades humanas, objetivam organizar, armazenar e possibilitar o acesso às informações neles existentes, contribuindo para a disseminação da informação e preservação da memória que, segundo Fragoso (2009), é um patrimônio cultural da sociedade e é responsável por alimentar a história da nação.

Nesse contexto, os sistemas que trabalham para tornar a recuperação possível são conhecidos como Sistemas de Recuperação de Informação, e têm como propósito tornar acessível o grande volume de documentos produzidos, de maneira que possam ser recuperados e disseminados, possibilitando a recuperação e o entendimento da memória produzida e armazenada neles. Lunardi (2008) destaca, ainda, que o principal objetivo de um SRI é a seleção, dentro

de um universo de documentos disponíveis, do conjunto de documentos relevantes para uma necessidade de informação do usuário. Nessa mesma linha de raciocínio, Araújo Junior (2007) afirma que os SRIs dizem respeito a um sistema de operações interligadas para identificar, dentre um grande conjunto de informações, aquelas que são, de fato, pertinentes, ou seja, que estão de acordo com a necessidade expressa pelo usuário.

É importante explicar, neste momento, que aqui se consideram os adjetivos ‘relevante’ e ‘pertinente’, anteriormente utilizados, como termos empregados para qualificar documentos úteis, ou seja, documentos recuperados que contribuem para atender a uma necessidade de informação. Dessa forma, nesta pesquisa, os termos ‘pertinente’, ‘útil’ e ‘relevante’ são utilizados como referência a documentos que foram recuperados e que são considerados adequados para satisfazer a necessidade informacional do usuário.

É certo que, para que os SRIs viabilizem o acesso aos itens de informação e recuperem documentos relevantes, precisam envolver três principais atividades: 1- Representação das informações contidas nos documentos, que é realizada, usualmente, através dos processos de indexação e descrição dos documentos; 2- Armazenamento e gestão física e/ou lógica desses documentos e de suas representações; 3- Recuperação das informações representadas e dos próprios documentos armazenados, de forma a satisfazer as necessidades de informação dos usuários (SOUZA, 2006).

Na primeira atividade, os assuntos tratados nos documentos são representados no sistema. Essa representação ocorre, geralmente, através do processo de indexação, a qual consiste na atribuição de termos (palavras-chave) a documentos, com a finalidade de representar as informações que estes contém, buscando descrever ou identificar cada documento através de seu conteúdo. Segundo Rowley (2002, p. 162), “indexação é o processo de atribuir termos ou códigos de indexação a um registro ou documento, termos ou códigos esses que serão úteis posteriormente na recuperação do documento ou registro”.

Lancaster (2004) aborda duas etapas principais que envolvem a indexação: análise conceitual e tradução. A análise conceitual visa identificar de que trata o documento, ou seja, qual o seu assunto. Deve, também, ser adaptada para se ajustar às necessidades de determinada clientela, ou determinado grupo de usuários. Assim, o objetivo nessa etapa é tentar antecipar quais termos seriam utilizados na busca, facilitando, dessa forma, o momento da recuperação da informação.

A tradução, a segunda etapa da indexação, envolve a conversão da análise conceitual de um documento para um determinado conjunto de termos de indexação, ou linguagem de indexação. A esse respeito, faz-se uma distinção entre indexação por extração (indexação derivada), que ocorre quando o conteúdo temático do documento é representado por palavras ou expressões que constam no próprio documento; e por atribuição, que consiste numa representação temática por meio de termos selecionados a

partir de uma fonte que não é o próprio documento. Os termos podem ser atribuídos pelo indexador ou de algum vocabulário controlado.

Outro aspecto importante acerca da indexação é a possibilidade de ser realizada de duas formas: intelectual ou automática. Na indexação intelectual, também chamada indexação manual, profissionais elegem termos capazes de representar as informações nos documentos, por intermédio do processo intelectual de análise de assunto. Na indexação automática, softwares atribuem termos julgados relevantes para representar as informações presentes nos documentos por meio de análise estatística da ocorrência de palavras, de funções lingüísticas de grupos de palavras, ou de diferentes métodos desenvolvidos para programas de computador. Nesse aspecto, Araújo Júnior (2007) descreve a indexação automática como sendo qualquer procedimento que permita identificar e selecionar os descritores que representam o assunto dos documentos, sem a intervenção humana direta.

A partir da abordagem das duas formas de realizar a indexação, ou seja, intelectual ou automática, pode-se questionar como identificar a indexação mais adequada. Para essa reflexão, Lancaster (2004, p. 83) aponta, de modo muito pragmático, que uma adequada indexação é aquela “que permite que se recuperem itens de uma base de dados durante buscas para as quais sejam respostas úteis, e que impede que sejam recuperados quando não são respostas úteis”. Por isso, seguindo o pensamento do autor supra citado, tanto a forma intelectual quanto a automática podem ser

consideradas adequadas, desde que auxiliem na recuperação de respostas úteis.

Vale salientar, ainda, que a atividade de indexação origina produtos que são conhecidos como índices. Como argumenta Feitosa (2006), os índices são instrumentos utilizados para a representação do conteúdo dos documentos, e têm como função facilitar a recuperação de informações relativas ao documento indexado ou resumido. O índice pode ser visto, ainda, como a estrutura que permite aos SRIs selecionarem documentos que satisfaçam uma busca; neste contexto, o índice geralmente consiste de uma estrutura de dados chamada arquivo invertido (BAEZA-YATES; RIBEIRO NETO, 1999). Nessa estrutura, as palavras apontam para os documentos onde ocorrem e, adicionalmente, podem ser armazenadas também a frequência de ocorrência da palavra e as posições em que ela ocorreu em cada documento. Assim, é possível fazer o cruzamento entre os registros e as solicitações do usuário, auxiliando, dessa forma, na atividade final do SRI, a recuperação.

O índice, produto final da indexação, revela importantes características da coleção de documentos e do SRI, como o vocabulário do sistema, as relações de coocorrência¹ das palavras e as relações de similaridade entre os documentos, porém, o índice deve permanecer em

¹ Revelam as palavras que ocorrem em combinação com outras na coleção de documentos. Assim, a partir de uma determinada palavra é possível evidenciar as palavras que coocorrem com ela e, também, indicar em que documentos elas podem ser encontradas. Inclusive, esse conceito foi trabalhado nos resultados da pesquisa.

constante processo de atualização para que possa auxiliar na precisão da recuperação.

A segunda atividade de um SRI é o armazenamento, que pode ser entendido como a forma como os SRIs guardam suas informações para, posteriormente, serem recuperadas; neste caso, utiliza-se o próprio computador tanto para armazenar os arquivos de documentos, como para manutenção das bases de dados (ROWLEY, 2002). Os recursos tecnológicos existentes permitem o aumento na capacidade de armazenamento dos computadores, o que possibilita que os sistemas de informação e bases de dados armazenem e disponibilizem um número cada vez maior de documentos.

A partir do armazenamento das informações no SRI, torna-se possível a realização das atividades de busca ou consulta, que consistem no momento em que o usuário expressa sua necessidade ao sistema com o propósito de recuperar documentos pertinentes ali armazenados. A respeito disso, Araújo Júnior (2007) comenta que o processo de busca para recuperação de informação pode ser conceituado como o processo de localizar documentos e itens de informação que tenham sido armazenados, com a finalidade de permitir o acesso dos usuários aos itens de informação, objetos de uma solicitação. Sendo assim, como afirma Gonzáles e Lima (2003), as buscas constituem a entrada de um sistema de recuperação de informação, ou seja, a busca se configura como o meio de o usuário se comunicar com o sistema para expressar a sua necessidade de informação.

Neste aspecto, tem-se **a terceira atividade**, a recuperação de informação, que inicia através da análise da necessidade expressa pelo usuário e redonda na apresentação de um conjunto ordenado de documentos, possivelmente permitindo ao usuário um *feedback* sobre aqueles que lhe atendem de uma maneira mais útil (SOUZA, 2006).

No entanto, os SRIs fornecem dois modos por meio dos quais o usuário pode recorrer ao sistema, em busca de itens de informação que atendam sua necessidade: através do *modo de recuperação* ou através do *modo de navegação*. Em outras palavras, quando o usuário recorre a um SRI, ele pode proceder de duas formas: expressando sua necessidade por meio da formulação de uma questão de busca ou, também, por meio de palavras-chave; nesta hipótese, tem-se o *modo de recuperação*; ou navegando pelo sistema e examinando os documentos em busca de informações pertinentes, processo conhecido como *browsing*, caso em que se verifica o *modo de navegação*.

No *modo de navegação*, o usuário navega através de páginas selecionando links em busca de documentos e informações que lhe sejam úteis (GONZÁLES; LIMA, 2003). Nesse modo, o usuário não tem uma necessidade expressa através de uma consulta, ele percorre o sistema, por meio do *browsing*, analisando os documentos que possam preencher suas lacunas informacionais. Desta forma, o usuário tem a possibilidade de localizar outros documentos que, anteriormente, não estavam no seu campo de interesse, mas

que podem vir a ser pertinentes. Baeza-Yates e Ribeiro Neto (1999) também concordam com esse aspecto, ao destacarem que o objetivo do usuário que realiza o *browsing* pode não estar claramente definido inicialmente, seus propósitos podem mudar durante a interação com o sistema.

De outra forma, quando o modo de busca ocorre por *recuperação*, o usuário de um sistema precisa traduzir a sua necessidade de informação por meio de uma expressão de busca, através de uma linguagem natural² ou através de uma linguagem artificial³, isto é, uma linguagem controlada. Geralmente a expressão de busca é composta por um conjunto de palavras que tentam exprimir a semântica da necessidade de informação do usuário, e resulta na recuperação de um número de documentos que possibilite a verificação de cada um deles, a fim de selecionar os que são pertinentes (FERNEDA, 2003).

Sobre isso, Rowley (2002) explica que o modo de recuperação envolve três etapas: 1- aceitação de uma consulta como uma representação da necessidade de informação, formulada pelo usuário; 2- execução de uma comparação da consulta com os registros existentes na base de dados (representações dos documentos); e por fim, 3- produção como resultado, a ser submetido ao usuário, de um

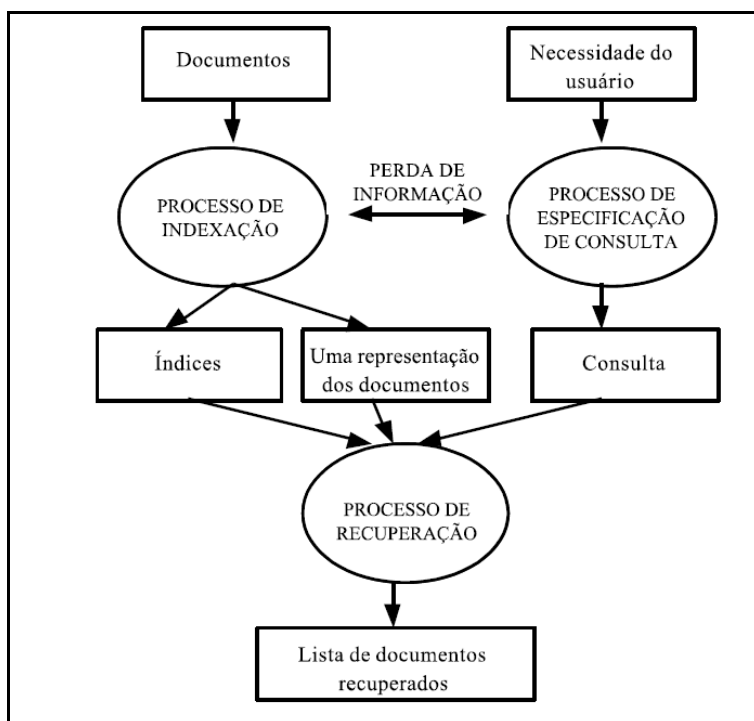
² Linguagem natural formada por termos livres que expressam o conteúdo dos documentos sem um controle terminológico específico. Consiste, também, em linguagem facilmente utilizada pelo usuário na busca por informações presentes em documentos.

³ Linguagem documentária, também denominada de linguagem de indexação. Consiste em uma linguagem artificial, formada por termos autorizados para o registro e representação dos assuntos contidos nos documentos. Essa linguagem é dotada de vocabulário controlado e regida por sintaxe própria (CAVALVANTI, 1978, apud CAFÉ; SALES, 2010).

conjunto de registros recuperados e que foram identificados com base nessa comparação.

Diante do exposto, buscando visualizar de forma mais clara como estão ordenadas as atividades em um sistema, Gey (1992 apud CARDOSO, 2000) apresenta uma ilustração, disposta na figura 1 a seguir, acerca de como pode estar estruturado um SRI e seus componentes.

Figura 1 – Componentes de um sistema de recuperação de informação



Fonte: Cardoso (2000, p. 1)

Na figura 1, são apresentados os componentes do sistema onde estão incluídos: documentos; a representação (indexação) dos documentos; necessidade do usuário expressa por meio da consulta (expressão de busca) e, finalmente, o processo de recuperação que, a partir dos documentos armazenados e da consulta formulada, recupera uma lista de documentos, considerados pelo sistema como sendo relevantes.

Portanto, a partir da análise das atividades que compõem um SRI, entende-se que a recuperação é a atividade final do processo, atividade em que se obtém o documento desejado. No entanto, a atividade de recuperar documentos que atendam, de forma satisfatória, a uma questão específica do usuário configura-se, ainda, como uma dificuldade para os SRIs. Essa dificuldade será melhor discutida a seguir.

2.1 Dificuldade enfrentada pelos Sistemas de Recuperação de Informação

Vale ressaltar que os SRIs são formados por três atividades que, como mencionado anteriormente, consistem em: representação, armazenamento e recuperação. Como afirma Ingwersen (2002), essas três atividades resultam em um objetivo final, o de recuperar documentos que contenham informações relevantes para uma questão formulada pelo usuário. No entanto, esse objetivo só será alcançado com

êxito se houver um bom desempenho das atividades que o envolvem.

Assim, a adequada realização de cada uma das atividades pode garantir que o máximo de itens relevantes seja recuperado e chegue até o usuário, auxiliando-o na satisfação de sua necessidade. Como enfatiza Rowley (2002), a questão crucial é que o processo de recuperação depende muito das etapas de representação (indexação) e armazenamento, as quais determinam, em grande medida, a melhor estratégia para as buscas feitas em um sistema de recuperação de informação.

De acordo com Araújo Júnior (2007), mais do que partes do sistema de recuperação de informação, o armazenamento e, sobretudo, a indexação, são os seus fatores críticos de sucesso, pois, segundo o autor, para que o armazenamento seja realizado com efetividade a indexação precisa ser feita de modo satisfatório, ou seja, representando adequadamente o conteúdo dos documentos. Com base no exposto pelo autor, percebe-se que a indexação, como meio de representação da informação, tem sua importância evidenciada. Tal fato é explicado pelo referido autor ao mencionar que a indexação influi diretamente na recuperação da informação, pois se os descritores selecionados para representar o conteúdo de um dado documento não forem coerentes, certamente o documento não será recuperado com facilidade, fato que comprometerá o processo como um todo.

Desta maneira, evidencia-se que a indexação exerce uma notória influência sobre a atividade de recuperação dos SRIs e que, relacionada à indexação está, também, a

dificuldade enfrentada por esses sistemas. Tal dificuldade reside em, diante de uma solicitação do usuário, recuperar documentos armazenados no acervo, que sejam potencialmente úteis dentro do contexto de seu interesse. Tendo em vista que, por um lado, os sistemas tendem a retornar uma grande quantidade de documentos, dentre os quais muitos se tornam irrelevantes e desnecessários ao usuário, ou, por outro lado, os sistemas retornam uma quantidade reduzida de documentos, ocultando, dessa forma, aqueles que poderiam ser úteis.

Essa dificuldade acarreta problemas nas respostas a questões específicas, como baixas taxas de revocação e precisão. Sobre o assunto, Souza (2006) explica que a “revocação”, ou “*recall*” ou mesmo “abrangência”, é a razão do número de documentos atinentes recuperados sobre o total de documentos atinentes disponíveis na base de dados, medindo, dessa maneira, o sucesso do SRI em recuperar documentos pertinentes. O autor entende por precisão a razão do número de documentos atinentes recuperados sobre o total de documentos recuperados, medindo, assim, o sucesso do SRI em não recuperar documentos que não sejam relevantes de acordo com a necessidade de informação.

Diante disso, o esperado é que os SRIs estabeleçam uma boa relação entre os índices de revocação e precisão, de forma a recuperar o maior número possível de documentos relevantes e o menor número possível de documentos pouco ou não relevantes, de acordo com as necessidades de informação do usuário.

Sendo assim, dois fatores que influenciam a revocação e a precisão e, conseqüentemente, o desempenho do SRI, são a exaustividade e a especificidade, fatores que são diretamente atribuíveis à indexação, isto é, à representação adotada pelo sistema. A especificidade e exaustividade, de acordo com Pinho (2010), estão intimamente ligadas ao momento da seleção dos conceitos candidatos à indexação, o que é feito após a análise do assunto do documento, porque tanto a primeira quanto a segunda, em maior ou menor intensidade, dependerão desta seleção.

Dessa maneira, uma das decisões importantes no processo de indexação é a definição sobre sua exaustividade. A exaustividade diz respeito à extensão ou números de conceitos escolhidos para representar o conteúdo de determinado documento (FEITOSA, 2006). Lancaster (2004, p. 27) argumenta, ainda, que uma indexação exaustiva “implica o emprego de termos em um número suficiente para abranger o conteúdo temático do documento de modo bastante completo. A indexação seletiva, por outro lado, implica o emprego de uma quantidade muito menor de termos, a fim de abranger somente o conteúdo temático principal do documento”.

Sendo assim, quanto mais termos forem utilizados para indexar um documento mais vezes ele será recuperado em uma busca, embora não exista uma quantidade definida para a atribuição de termos. Dessa forma, a escolha de muitos termos para representar o conteúdo do documento, isto é, a exaustividade, poderá acarretar no aumento da revocação no momento da busca.

A especificidade, por sua vez, diz respeito “à exatidão com que os termos escolhidos descrevem o conteúdo do documento em questão” (FEITOSA, 2006, p. 29). A especificidade é considerada por Lancaster (2004) o mais importante princípio da indexação, já que, segundo ele, um documento deveria ser indexado sob o termo mais específico que o abranja completamente.

Pode-se apresentar, como exemplo, um documento que trate do cultivo de laranjas; no caso da indexação que adote a especificidade, o documento seria indexado sob “laranjas” e não sob “frutas cítricas” ou “frutas”. Dessa forma, esse princípio busca atribuir à indexação termos capazes de especificar cada documento, de forma que sua recuperação seja mais precisa, evitando a atribuição de termos genéricos que retornem para o usuário uma série de outros documentos relacionados ao assunto, porém em categorias mais amplas. Diante disso, a escolha de poucos termos, mas específicos, para representar o conteúdo de um documento – a especificidade – poderá acarretar o aumento da precisão no momento da busca.

Baseando-se nas abordagens de Lancaster (2004) acerca dos princípios da indexação e relacionando-os com os índices de revocação e precisão, Medeiros (2010) desenvolveu o quadro a seguir.

Quadro 1 - Princípios da indexação

Princípio	Definição	Aumenta	Diminui
Exaustividade	Corresponde, de maneira geral, ao número de termos atribuídos, mas não há relação unívoca entre exaustividade e número de termos atribuídos.	Revocação	Precisão
Especificidade	Um tópico deve ser indexado sob o termo mais específico que o abranja completamente.	Precisão	Revocação

Fonte: Medeiros (2010)

Com base no quadro 1, pode-se concluir que a exaustividade prevê um número maior de termos para os documentos, aumenta a possibilidade de recuperação e aumenta a revocação; no entanto, diminui a precisão, ou seja, será recuperado um número maior de itens, contudo, muitos deles poderão tratar apenas superficialmente do assunto, não sendo relevante para a necessidade do usuário. Por outro lado, quanto maior a especificidade, maior será a probabilidade de alcançar alta precisão. Porém, a recuperação de documentos específicos acarreta a diminuição da revocação, ou seja, diminuição na recuperação de outros documentos presentes no sistema, relacionados ao assunto desejado.

A esse respeito, Feitosa (2006), baseando-se em Lancaster (2003), expõe sua opinião ao afirmar que o ideal é que se aumente o nível de precisão, isto é, a capacidade de evitar a recuperação de documentos inúteis, por meio do aumento da especificidade da indexação, pois quanto maior for o acervo da base de dados do sistema, menos aceitável será uma baixa precisão. Nessa linha de pensamento, Lancaster (2003, p. 31) assinala que:

à medida que as bases de dados crescem de tamanho, a quantidade de itens que aparecem sob qualquer termo também tende a crescer. Tornando-se necessário, portanto, indexar com o emprego de mais termos (e também torná-los cada vez mais específicos) de modo que a indexação seja mais discriminativa para possibilitar pesquisas em que se alcance um nível adequado de revocação com nível tolerável de precisão.

A preocupação com uma indexação precisa e pertinente se deve ao fato de que um sistema de informação deve sempre ser organizado com vistas à recuperação eficiente da informação. Para isso, a exaustividade e a especificidade exercem importante papel no resultado da indexação, refletindo diretamente na precisão e revocação, ou seja, na capacidade de recuperar documentos pertinentes e de evitar que documentos não pertinentes sejam recuperados, resultando, dessa forma, na recuperação de informações que sejam úteis à necessidade do usuário e, conseqüentemente, contribuindo para o sucesso do SRI, pois, como argumenta Feitosa (2006, p. 28), “resultados satisfatórios na recuperação dependem diretamente da qualidade com que a indexação foi realizada [...]”.

Sendo assim, após a realização de uma adequada indexação, e do armazenamento daquilo que foi indexado, tem-se a terceira atividade: a recuperação propriamente dita, atividade em que o resultado de todo o empenho empregado anteriormente será evidenciado. Para isso, ou seja, para a realização da terceira atividade, os SRIs fazem uso de modelos específicos, conhecidos como modelos de recuperação de informação. A esse respeito, Ferneda (2003, p.18) destaca que “a eficiência de um sistema de recuperação de informação está diretamente ligada ao modelo que o mesmo utiliza”. Para isso dedica-se uma seção específica para os modelos, através dos quais os SRIs realizam a localização e recuperação da informação que o usuário deseja.

2.2 Modelos de Recuperação de Informação utilizados em SRIs

Serão destacados aqui os modelos de recuperação de informação que se enquadram na terceira atividade de um SRI, e estão relacionados ao modo como o sistema procede para atender ao usuário que recorre a ele em busca de documentos que satisfaçam sua necessidade.

Diante da existência de diversos modelos de recuperação de informação, serão adotados, para a presente pesquisa, aqueles abordados por Ferneda (2003), que estão divididos em modelos quantitativos e modelos dinâmicos. Apesar de se julgar relevante a exposição dessas duas categorias de modelos, vale salientar que aqueles que mais

interessam ao propósito da pesquisa em questão são os modelos quantitativos.

A justificativa para a preferência pelos modelos quantitativos reside no fato de que eles, por serem modelos direcionados aos sistemas automatizados e de funcionalidades comprovadas, são os que mais se adéquam ao objetivo proposto nesta pesquisa, qual seja, o de relacionar a organização da informação e a visualização da informação, analisando a contribuição que a visualização tem na construção de interfaces gráficas para sistemas de recuperação de informação. Vale ressaltar, ainda, que os modelos dinâmicos se destacam como propostas mais recentes, que utilizam métodos da inteligência artificial e que se configuram como alternativas a serem estudadas futuramente; além disso, o aprofundamento nesses modelos ampliaria por demais o espectro da pesquisa, inviabilizando sua condução em tempo adequado.

Nesse contexto, serão abordados a seguir os modelos expostos por Ferneda (2003).

2.2.1 Modelos Quantitativos

Os clássicos modelos quantitativos de recuperação de informação abrangem: modelo booleano, modelo vetorial, modelo probabilístico, modelo *fuzzy*, e modelo booleano estendido. A grande maioria dos modelos de recuperação de informação, segundo Ferneda (2003), é de natureza quantitativa, baseados em disciplinas como a lógica, a

estatística e a teoria dos conjuntos. Nos modelos quantitativos, que são muito utilizados por SRIs e mecanismos de busca na web, os documentos são representados por um conjunto de termos de indexação, ou seja, por palavras que buscam representar o assunto dos documentos.

O Modelo Booleano é baseado na teoria dos conjuntos e consiste em um sistema binário, no qual existem somente dois valores possíveis para qualquer símbolo algébrico, ou seja, 1 ou 0, verdadeiro ou falso. Para cada questão de busca, são recuperados todos os documentos que possuem os termos nas condições especificadas pelo usuário. As buscas são formuladas através de uma expressão booleana composta por termos ligados através dos operadores lógicos AND, OR e NOT (E, OU e NÃO). Por isso, os operadores booleanos estabelecem relações específicas de ocorrência com as palavras-chave, de forma a especificar os documentos a serem recuperados.

Dessa maneira, uma expressão conjuntiva de enunciado T1 **AND** T2 recuperará documentos indexados por ambos os termos (T1 e T2). Uma expressão disjuntiva T1 **OR** T2 recuperará o conjunto dos documentos indexados pelo termo T1 ou pelo termo T2. Uma expressão que utiliza apenas um termo T1 terá como resultado o conjunto de documentos indexados por T1. As expressões T1 **NOT** T2 ou T1 **AND NOT** T2 terão o mesmo resultado, o conjunto dos documentos indexados por T1 e que não são indexados por T2. Nesse caso o operador **NOT** pode ser visto com um operador da diferença entre conjuntos. Assim, termos e operadores podem

ser combinados para especificar buscas mais detalhadas ou restritivas (FERNEDA, 2003).

A maior desvantagem do modelo booleano, segundo Souza (2006), é o fato de trabalhar de forma binária, ou seja, os documentos são analisados sob o critério dualista relevante/não relevante; dessa forma, as respostas às questões de busca podem ser nulas ou muito grandes, como bem destaca Cardoso (2010). Outra desvantagem consiste no fato de não ser criada nenhuma espécie de ordenação dos resultados que atendam às condições da busca.

Apesar das desvantagens, muitos sistemas se desenvolveram utilizando o modelo booleano como ponto de partida para implementação de novos recursos de recuperação (FERNEDA, 2003).

O Modelo vetorial propõe um ambiente no qual os documentos respondam parcialmente a uma expressão de busca. Neste modelo, pode-se calcular o grau de similaridade entre os documentos em relação à busca, através da associação de pesos tanto aos termos de indexação como aos termos da expressão de busca formulada pelo usuário.

Sobre esse aspecto, Ferneda (2003) explica que, no modelo vetorial, um documento é representado por um vetor em que cada elemento representa o peso ou a relevância do respectivo termo de indexação para o documento. Cada vetor descreve a posição do documento em um espaço multidimensional, onde cada termo de indexação representa uma dimensão ou eixo. Cada elemento do vetor (peso) é normalizado de forma a assumir valores entre zero e um. Os

pesos mais próximos de um (1) indicam termos com maior importância para a descrição do documento.

Uma considerável limitação desse modelo é não permitir a formulação de buscas booleanas, o que restringe consideravelmente sua flexibilidade. Acerca disso, Gonzáles e Lima (2003) apontam as dificuldades enfrentadas por esse modelo em explicar relacionamentos específicos de sinonímia, e também relacionamentos entre termos compostos por mais de uma palavra, tarefas que seriam facilmente realizáveis através da utilização de operadores lógicos como OU e E.

No entanto, apesar de tudo, de acordo com Souza (2006, p. 167), “o modelo vetorial é a base da grande maioria de sistemas de recuperação de informações, mais notadamente os que têm como objeto a Internet [...]”. Esse fato pode ser decorrente das vantagens apontadas por Cardoso (2010) acerca do modelo, como a simplicidade, facilidade de se computar similaridades com eficiência, e por se comportar bem com coleções genéricas.

O Modelo probabilístico procura representar o processo de recuperação de informação sob um ponto de vista probabilístico e, para isso, supõe-se que exista um conjunto ideal de documentos que satisfaça a cada uma das consultas ao sistema, ou seja, os documentos relevantes para o usuário, e que este conjunto possa ser recuperado. Para recuperação, o modelo probabilístico considera um processo iterativo de estimativas da probabilidade de relevância (CARDOSO, 2010).

Assim, através de tentativa inicial, quando o usuário formula uma expressão de busca, o modelo trabalha com a

probabilidade do termo da questão estar presente em um conjunto de documentos e de ser recuperado como item importante. A partir da ordenação do conjunto de documentos recuperados na primeira busca, o usuário seleciona alguns documentos que considera relevante para sua necessidade, então, o sistema utiliza essa informação e tenta melhorar os resultados subsequentes.

Através do *feedback* do usuário em sucessivas interações, busca-se aproximar cada vez mais do conjunto ideal, por meio de análise dos documentos considerados pertinentes pelo usuário.

O valor desse modelo está em considerar a interação com o usuário como um caminho para refinar o resultado continuamente (SOUZA, 2006). Por isso, sua maior virtude é o fato de reconhecer que a atribuição de relevância é uma tarefa do usuário; sendo assim, o modelo probabilístico incorpora explicitamente o processo de *Relevance Feedback*⁴ como base para a sua operacionalização.

No Modelo *fuzzy* busca-se estender o conceito da representação dos documentos por palavras-chave, assumindo que cada questão de busca determina um

⁴ *Relevance feedback* (realimentação por relevância) é a mais popular estratégia de reformulação de questões de busca. Em um ciclo de *relevance feedback*, uma lista com os documentos recuperados é apresentada ao usuário que, depois de examiná-los, seleciona aqueles que são relevantes. A principal ideia consiste na seleção de termos importantes, ou expressões, anexadas ao documento que foram identificados como relevante pelo usuário, salientando a importância desses termos em uma nova formulação de questão de busca. O efeito esperado é que a nova questão seja direcionada aos documentos relevantes e afastada daqueles não relevantes.

conjunto difuso e que cada documento possui um grau de pertencimento a esse conjunto (SOUZA, 2006).

No modelo *fuzzy* um documento pode ser visto como um conjunto fuzzy⁵ de termos, cujos pesos dependem do documento e do termo em questão. O peso associado ao termo expressa o quanto esse termo é significativo na descrição do conteúdo de um documento. A qualidade da recuperação pode, então, ser influenciada pela função adotada para calcular os pesos dos termos de indexação. Dessa forma, um documento possui um grau de pertencimento a uma questão de busca, de acordo com o peso estabelecido para os termos que o representam.

Assim, para cada *tag* pode ser associado um valor numérico que expressa a sua importância para o documento. O peso de um termo em relação a um determinado documento é obtido através de uma função de agregação, que considera a importância de cada *tag* do documento onde o termo aparece. A teoria dos conjuntos *fuzzy* trabalha com um grau de imprecisão, ou seja, com a medida de possibilidade de um elemento pertencer ou ser membro de um dado conjunto (FERNEDA, 2003).

⁵ O termo *fuzzy* significa vago, impreciso, nebuloso. A teoria de conjuntos *fuzzy* lida com a representação de classes ou conjuntos de elementos cujos limites não estão bem definidos. A ideia chave é associar uma função de relação entre os elementos da classe. Essa função toma valores no intervalo entre 0 e 1, sendo 0 correspondente a nenhuma relação de pertencimento à classe, e 1 correspondente a elementos pertencentes à classe (BAEZA-YATES; RIBEIRO NETO, 1999). Assim, a participação em um conjunto fuzzy é uma noção intrinsecamente aproximada, em vez de fixa e exata, ou seja, lida com o conceito de verdade parcial, em que o valor de verdade pode variar entre completamente verdadeira e completamente falsa.

Baeza-Yates e Ribeiro Neto (1999) destacam, ainda, a possibilidade da adoção de instrumentos como o tesauro para definir relacionamentos entre termos. Para os autores, a ideia básica é expandir o conjunto de termos na questão de busca com termos relacionados (obtidos do tesauro), que poderão influenciar na recuperação de documentos relevantes adicionais para a questão de busca, além dos que seriam normalmente recuperados. Dessa forma, o tesauro pode sugerir termos relacionados semanticamente que também confirmem algum grau de pertencimento ao conjunto difuso determinado pela questão de busca.

O Modelo booleano estendido procura unir a potencialidade das expressões booleanas com a precisão das buscas do modelo vetorial, tentando, dessa forma, contornar as limitações desses clássicos modelos: booleano e vetorial.

Por um lado, busca-se flexibilizar o modelo booleano, introduzindo uma gradação no conceito de relevância, pois, apesar de permitir uma recuperação relativamente eficiente, este modelo não possibilita o ordenamento dos documentos recuperados; bem como, por outro lado, dar maior poder às buscas do modelo vetorial através do uso dos operadores booleanos, tendo em vista que, apesar de permitir a ordenação dos documentos resultantes de forma precisa, não possibilita a utilização de buscas booleanas, o que restringe sua capacidade.

Como afirma Ferneda (2003), o modelo booleano estendido tenta minimizar as limitações do modelo vetorial e booleano clássicos através de uma conceituação matemática mais genérica. Essa generalização é feita através da

informação de dois novos parâmetros em relação ao modelo booleano tradicional: os pesos associados aos termos da expressão de busca e o parâmetro associado a cada operador booleano. Ainda segundo o autor, esse aumento da complexidade na formulação de buscas torna-se, também, a principal desvantagem do modelo booleano estendido.

Este modelo não tem sido utilizado extensivamente. No entanto, segundo Baeza-Yates e Ribeiro Neto (1999), o modelo possui propostas claras e que podem revelar-se útil no futuro.

2.2.2 Modelos Dinâmicos

Os modelos dinâmicos de recuperação de informação, abordados por Ferneda (2003), têm como principal característica o reconhecimento da importância da interação e interferência dos usuários na definição das representações dos documentos, proporcionando uma evolução ou adaptação dos documentos aos interesses dos usuários do sistema, percebidos por meio de suas buscas e da atribuição de relevância e não relevância aos documentos recuperados (*relevance feedback*).

Os modelos dinâmicos abrangem: sistemas especialistas; redes neurais e algoritmos genéticos. Tais modelos resumem propostas mais recentes que utilizam métodos derivados da Inteligência Artificial e que representam alternativas promissoras a serem estudadas e desenvolvidas futuramente.

Os Sistemas Especialistas procuram simular o conhecimento humano em um domínio específico, a partir do desenvolvimento de ferramentas computacionais, com o propósito de auxiliar na tomada de decisões e na resolução de problemas relacionados a esse domínio. Tendo em vista que a eficácia do processo de recuperação de informação depende em grande parte do conhecimento, a incorporação do conhecimento a um sistema poderia auxiliar nesse processo.

Um sistema especialista é composto principalmente de:

1- Base de conhecimento, que consiste no elemento central do sistema especialista, e é onde está representado e armazenado o conhecimento relevante sobre o problema. O conhecimento é, geralmente, representado por regras do tipo condição – ação, na forma Se - Então, **SE** (condição) **ENTÃO** (ação). A seguir apresenta-se um exemplo do sistema especialista MYCIN, desenvolvido durante a década de 70, com a finalidade de prescrever medicação para pacientes com infecções bacterianas (FERNEDA, 2003):

SE o paciente apresenta **febre**,
o paciente apresenta **vômitos** e **diarreia**,
o paciente está **desidratado**

ENTÃO o paciente sofre de **infecção intestinal**

SE o paciente sofre de **infecção intestinal**

ENTÃO o paciente deve tomar **ampicilina**.

2- Motor de inferência, que consiste no conjunto de métodos e instrumentos para manipular o conhecimento

armazenado na base, selecionando-o e aplicando-o na resolução do problema em questão. Esses instrumentos estão relacionados com a forma de inferência sobre os dados do sistema, com a forma como as regras da base de conhecimento serão testadas e com os métodos de tratamento e incerteza. A máquina de inferência busca as regras necessárias a serem avaliadas, ordena-as de maneira lógica e direciona o processo de inferência, baseado nos dados simbólicos contidos na base de conhecimento (FERNEDA, 2003).

Os exemplos de formas de representação de conhecimento em sistemas especialistas são as redes semânticas e os *frames*. A rede semântica é composta por um conjunto de nós que, em geral, representam objetos e que são conectados por um conjunto de arcos, os quais representam as relações existentes entre eles. Já o *frame* consiste em uma coleção de atributos e valores associados. Cada *frame* representa uma classe ou uma instância (elemento de uma classe). A criação de um sistema de *frames* é possível graças ao fato que o valor de um atributo de um *frame* pode ser outro *frame*. Dessa maneira, um sistema de *frames* pode definir uma hierarquia de classes (FERNEDA, 2003).

As Redes Neurais são modelos que buscam simular o processamento de informação do cérebro humano. São compostas por unidades de processamentos simples, os neurônios, que se unem através de conexões, formando uma rede neural. Souza (2006) explica que, neste modelo, utiliza-se o poder das redes neurais para realizar o casamento de padrões entre as questões de busca e os documentos do

acervo do sistema. Cada questão de busca “dispara” um sinal que ativa os termos índices que, por sua vez, propagam os sinais aos documentos relacionados. Esses, por sua vez, retornam os sinais a novos termos índices, em interações sucessivas. O conjunto resposta é definido através desse processo, e pode conter documentos que não compartilhem nenhum termo-índice com a busca, mas que tenham sido ativados durante o processo. Ferneda (2003, p. 66) explicita o processo de aplicação das redes neurais na recuperação de informação afirmando que:

uma tarefa comum para um sistema de recuperação de informação é pesquisar documentos relevantes que satisfazem uma determinada expressão de busca, através dos termos de indexação. Pode-se dizer que, em um sistema de recuperação de informação, de um lado, estão as expressões de busca, do outro lado estão os documentos e, no meio, ficam os termos de indexação. Essa estrutura pode ser vista como uma rede neural de três camadas: a camada de busca seria a camada de entrada da rede neural, a camada de documentos seria a saída e a camada de termos de indexação seria uma camada central.

As redes neurais, enquanto modelos de recuperação de informação, introduzem nos sistemas a capacidade de se adaptarem ao meio ambiente, ou seja, às buscas dos usuários.

No entanto, como aborda Baeza-Yates e Ribeiro Neto (1999), não há evidências conclusivas de que as redes neurais fornecem uma boa performance na recuperação de documentos em coleções mais gerais, isto é, o modelo não foi testado extensivamente em coleções com grande quantidade

de documentos. Ainda assim, os autores caracterizam as redes neurais como um paradigma de modelo alternativo, e destacam sua atraente funcionalidade: o fato de permitir a recuperação de documentos que não estão inicialmente relacionados com os termos da questão de busca.

Os Algoritmos genéticos baseiam-se na característica evolutiva para abordar o problema da recuperação de informação. A aplicação dos algoritmos genéticos na recuperação de informação representa um modelo inovador para todo o processo de recuperação.

Neste modelo, as representações dos documentos podem ser vistas como um tipo de “código genético”, no qual um cromossomo é representado por um vetor binário em que cada elemento armazena o valor 0 ou o valor 1, correspondendo, respectivamente, à ausência ou à presença de um determinado termo na representação do documento, identificando a relevância e a não relevância dos documentos para a questão de busca.

Durante o processo da busca, quando o usuário seleciona os documentos relevantes a sua necessidade de informação, processo também conhecido como *relevance feedback*, para um documento considerado relevante, as descrições responsáveis pela sua recuperação recebem um crédito pelo seu sucesso e as descrições que não participaram da sua recuperação são rebaixadas. Já para um documento recuperado considerado não relevante, as descrições responsáveis pela sua recuperação são rebaixadas e as demais descrições recebem um crédito.

Os sistemas de recuperação baseados nos algoritmos genéticos possuem a capacidade de evoluírem, alterando progressivamente as representações (código genético) dos documentos. Mas, apesar da característica evolutiva representar uma forma inovadora de abordar o problema da recuperação de informação, diversos questionamentos relacionados aos efeitos de sua inerente imprevisibilidade surgem quando utilizado em situações reais (FERNEDA, 2003).

Fazendo uma análise dos modelos de recuperação de informação aqui expostos, percebe-se que os modelos dinâmicos, diferentemente da maior parte dos modelos quantitativos, possibilitam a participação ativa dos usuários de um sistema de informação na representação dos documentos, porém, abrem um campo de discussões sobre sua aplicabilidade em circunstâncias reais. Assim, a utilização dos modelos dinâmicos em grandes comunidades de usuários, com interesses variados, poderia resultar em um efeito negativo, eliminando a principal vantagem desses modelos.

Por outro lado, apesar do aparente esgotamento dos modelos quantitativos, sua presença ainda é notada em sistemas de recuperação de informação, pois, como afirma Ferneda (2003), tais modelos ganharam força com os mecanismos de busca da web, que introduziram características específicas para tratar a quantidade de informação disponível na internet.

Diante do exposto acerca da recuperação da informação, fica claro que o uso da informação para satisfação de uma necessidade consiste no objetivo final de

um sistema de recuperação de informação, que tem como função permitir que o usuário localize o maior número possível de itens relevantes que se enquadrem a sua necessidade. Para isso, tem-se como condição primordial, como aponta Guedes (1994), a representação precisa do conteúdo temático dos documentos, que consiste no ponto vulnerável da recuperação da informação, como acrescenta Maculan, Lima e Penido (2011). Logo, pode-se perceber que, para que um sistema permita uma recuperação adequada que satisfaça o usuário com aquilo que ele deseja, se faz necessário uma atenção maior à maneira como as informações estão representadas e, também, organizadas no sistema.

Nessa mesma linha, analisando a abordagem feita por Esteban Navarro e García Marco (1995), a organização pode ser vista como um elemento mediador entre a representação e a comunicação; por um lado, a representação se efetua com a finalidade de permitir uma eficaz organização e, por outro lado, a comunicação exige uma correta recuperação, cujo êxito depende da qualidade da organização.

Neste sentido, recuperar informações relevantes que atendam à necessidade do usuário é a função dos SRIs. Para que isso seja possível, se faz necessário um olhar especial para a organização prévia dessas informações, pois, quando se trata de recuperação e de sistemas que lidam com informações e documentos, não se pode deixar de analisar os aspectos inerentes à organização da informação.

3 ORGANIZAÇÃO DA INFORMAÇÃO

O verbo “organizar”, de acordo com o Dicionário Houaiss, significa “dar determinada ordem a” (série de itens); “dispor de forma ordenada”; “dar forma regular às partes de” (um todo); “arrumar, ordenar”. Assim também se comporta o ser humano, organizando naturalmente os muitos tipos de coisas em sua vida: os livros em estantes, as roupas em armários, utensílios de cozinha em gavetas, documentos específicos em pastas etc. Essa organização, mesmo que inconscientemente, visa à facilitação de uma futura localização e recuperação dos itens armazenados.

De maneira geral, pode-se dizer que se organiza para poder compreender o mundo e se comunicar melhor. Por exemplo, ao explicar a alguém o significado de determinada palavra, normalmente recorre-se a sua classificação, incluindo a palavra em um grupo maior que a define. Assim, ao explicar o significado da palavra “cadeira”, seria dito que é um móvel (classe) que possui um assento e um encosto. Nesse momento, organiza-se no pensamento o conceito dessa palavra, focando a classe dos móveis e excluindo outras classes (CAFÉ; SALES, 2010).

No contexto dos sistemas de informação não é diferente. Os sistemas organizam um acervo, cuja função é armazenar objetos informacionais, compreendendo melhor de que tratam, dispondo de forma ordenada, de maneira que as informações neles registradas possam ser recuperadas com facilidade. Essa prática de armazenar e organizar o conteúdo informacional produzido, visando alcançar usuários futuros,

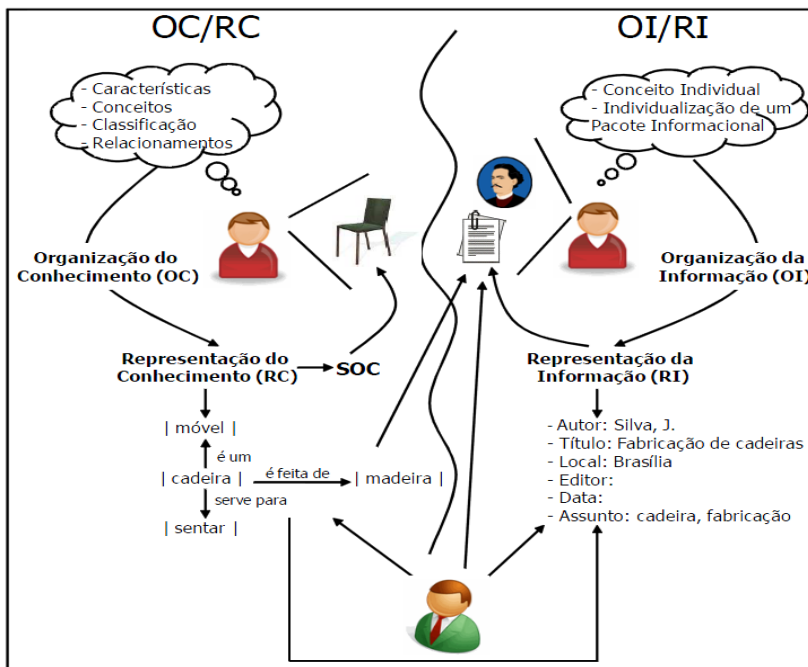
não é recente, como salienta McGarry (1999, p. 112): “[...] o anseio de armazenar os produtos da memória coletiva tem estado conosco desde o início da era da escrita”.

Antes de prosseguir com a abordagem acerca do conceito de Organização da Informação, seu significado e práticas, é importante diferenciá-lo do conceito de Organização do Conhecimento. Assim, Brascher e Café (2008), que discorrem em seu trabalho sobre a divergência entre os conceitos, abordam a existência de uma discordância entre a delimitação de Organização da Informação (OI) e Organização do Conhecimento (OC) que, por vezes, são utilizados com falta de clareza em seus sentidos.

Para as autoras, anteriormente citadas, a OI se aplica às ocorrências individuais de objetos informacionais, compreendendo a organização desses objetos para arranjá-los sistematicamente em coleções. A OC, por sua vez, se aplica às unidades do pensamento (conceitos), e visa a construção de modelos de mundo que se constituem em abstrações da realidade, que permitem descrever e fornecer explicações sobre os fenômenos que podem ser observados.

De forma resumida, a OI tem como objeto a informação registrada passível de ser armazenada e organizada, possibilitando sua recuperação quando necessário. Já a OC se encontra no mundo dos conceitos que visam à construção de representações do conhecimento. A figura 2, a seguir, ilustra essa proposta em relação a diferenciação entre os dois conceitos.

Figura 2 – OC/RC e OI/RI



Fonte: Brascher e Café (2008, p.7)

Com base no exposto acerca da OI e da OC, optou-se aqui pela utilização do termo Organização da Informação, já que sua proposta se adequa ao objetivo que se pretende alcançar com esta pesquisa, ou seja, analisar os aspectos da OI e sua influência na prática da organização que resulta na facilitação do acesso e recuperação de documentos informacionais em SRIs.

Dessa forma, considera-se a proposta conceitual de Brascher e Café (2008), em que a OI compreende a organização de um conjunto de objetos informacionais para arranjá-los sistematicamente em coleções, como a

organização da informação em bibliotecas, museus, arquivos, tanto tradicionais quanto eletrônicos. Considera-se, também, a contribuição de Alvarenga (2003), que acrescenta ainda que os processos de organização da informação destinam-se, prioritariamente, à recuperação eficaz de informações por parte dos usuários.

Assim, para melhor compreender como se desenvolvem as práticas da OI, se faz necessário uma análise mais aprofundada a respeito de um aspecto essencial que fundamenta suas atividades - o tratamento da informação.

3.1 Tratamento da informação

Toda informação produzida visa à sua efetiva utilização; com base nisso, tem-se na atividade de organização da informação um papel fundamental de caráter mediador, já que essa atividade é responsável pela comunicação entre a representação e o uso de informações. Ao entender que a informação só terá sentido se for transmitida e socializada, torna-se claro que, para que isso ocorra, faz-se necessário que ela circule e, para tanto, ela deve ser tratada nos seus aspectos formais e temáticos de forma adequada.

Como mencionado, para que a informação seja socializada, ela deve ser tratada de forma adequada. Contextualizando o tratamento da informação em sistemas de informação, Dias (2001, p.3) o define como “a função de descrever os documentos, tanto do ponto de vista físico (características físicas dos documentos) quanto do ponto de

vista temático (ou de descrição do conteúdo)”. Sendo assim, o tratamento da informação busca representar os documentos, através da síntese deles, para que sua organização e manipulação dentro do sistema de recuperação possam ser facilitadas e para que o usuário possa avaliar, de maneira mais simples, a relevância que os documentos apresentam para sua necessidade de informação.

Nessa linha de pensamento, Foskett (1973 apud GUIMARÃES, 2009) reconhece que a OI é uma atividade de natureza eminentemente mediadora, em que dois universos se descortinam: o primeiro, ligado ao acesso físico aos documentos; e o segundo, de natureza mais complexa, voltado para o acesso ao conteúdo informacional. Desse modo, a organização da informação é composta por dois grandes processos que estão relacionados entre si: o tratamento descritivo e o tratamento temático.

3.1.1 Tratamento Descritivo da Informação

O Tratamento Descritivo da Informação (TDI) consiste na distinção das características físicas de um documento, de modo que, a partir da sua representação, tal documento possa ser acessado, atingindo, assim, um objetivo maior: o de atender à demanda do usuário (MEY, 1995).

O tratamento descritivo é voltado aos elementos relativos à forma dos documentos, com base na análise do seu tipo e identificação das informações descritivas (como na catalogação descritiva). O TDI compreende um exame da informação a ser organizada, com o objetivo de identificar

alguns elementos nela contidos, os quais servirão para sua posterior identificação no momento da recuperação. Nesse processo, Nonato (2009) destaca que se torna necessária a utilização de instrumentos para tratamento descritivo que padronizam a descrição, podendo auxiliar no estabelecimento e cumprimento de certas regras, que possibilitam a correta identificação da informação manifestada na forma de um determinado documento (tais instrumentos serão abordados na subseção 3.1.1.2).

Como mencionado, o TDI dá-se por meio da catalogação, que consiste na representação do item com base em suas características extrínsecas. A partir da visão de Garrido Arilla (2001), pode-se entender que o propósito da catalogação tem sido, historicamente, a criação e organização da informação bibliográfica para proporcionar acesso às coleções de documentos em bibliotecas ou em sistemas de informação.

Nessa linha de pensamento, a catalogação é vista por Mey (1995, p.5) como:

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

Assim, o TDI, realizado através da catalogação, resulta em um registro que, em sua estrutura, engloba três conjuntos de elementos: a descrição bibliográfica, os pontos de acesso e os dados de localização (INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS, 2005; MEY, 1995).

A **descrição bibliográfica** é responsável pela representação sintética e codificada das características de um documento, de forma a individualizá-lo, ou seja, tornando-o único entre os demais. Para a descrição bibliográfica de um documento, as informações a serem descritas enquadram-se em oito áreas: título e responsabilidade, edição, detalhes específicos do material, dados de publicação, descrição física, série, notas e número internacional normalizado do item. Os **pontos de acesso**, por sua vez, consistem na parte pela qual os usuários podem acessar a representação de um item no catálogo, e podem ser definidos pelos termos que representam um dado documento, como nome, título, local. Já os **dados de localização** estão voltados para a identificação do documento; esses dados constituem-se em códigos que permitem ao usuário localizar o documento em determinado acervo, podendo compreender: código da biblioteca e número de chamada (MEY, 1995).

Os três conjuntos de elementos acima citados são responsáveis, respectivamente, pela individualização de diferentes manifestações de uma obra, pela reunião de suas semelhanças, e pela obtenção do documento, através da sua localização em determinado acervo (ZAFALON, 2012).

Entende-se, ainda, que a catalogação possui funções a serem desempenhadas, e que o desempenho de tais funções, como destacado por Mey (1995), se torna possível quando o processo de catalogação utiliza regras e padrões para a construção de catálogo como canal de comunicação de suas mensagens. Nesse contexto, baseado na autora, tem-se como funções da catalogação:

- a) Permitir ao usuário: 1- localizar um item específico; 2 - escolher entre várias manifestações de um item; 3 - escolher entre vários itens semelhantes, sobre os quais, inclusive, possa não ter conhecimento prévio algum; 4 - expressar, organizar ou alterar sua mensagem interna.
- b) Permitir a um item encontrar seu usuário⁶.
- c) Permitir a outra biblioteca: 1 - localizar um item específico; 2 - saber quais os itens existentes em acervos que não o seu próprio.

Vale salientar que, além de permitir a representação de um documento e sua localização, a catalogação permite, também, por meio da padronização, o intercâmbio dos registros. A esse respeito, Corrêa (2008) destaca que a padronização na representação das informações de documentos tem importante valor e é, também, o objeto de trabalho da catalogação como forma de garantia do intercâmbio de registros bibliográficos.

Convém ainda enfatizar que, para o desempenho das funções da catalogação, enquanto processo para construção de formas de representação descritiva, alguns aspectos precisam ser resguardados, como integridade, clareza, precisão, lógica e consistência, tendo em vista, constantemente, o usuário e as necessidades por ele demonstradas, requeridas ou determinadas (MEY, 1995; ZAFALON, 2012).

⁶ Um catálogo adequado torna possível ao item (objeto informacional) ir de encontro ao seu usuário. Isso é possível, já que, por meio de uma representação adequada, um item pode despertar o interesse do usuário, mesmo que, inicialmente, esse item não esteja dentro do seu campo de interesse. Nesse sentido, pode-se afirmar que o item encontrou o seu usuário.

No intuito de melhor compreender os fatos que marcaram a trajetória do tratamento descritivo da informação, apresenta-se a seguir uma breve fundamentação teórica acerca da catalogação, tendo em vista que é através dela que se configura o TDI.

3.1.1.1 Fundamentos teóricos

Nesta subseção, busca-se destacar os acontecimentos que nortearam e norteiam os processos referentes à prática da catalogação, ou seja, os aspectos que refletem o TDI. Aqui se restringe a destacar, de forma resumida, alguns dos notórios acontecimentos que exerceram influência na organização da informação, no âmbito do tratamento descritivo, tendo como foco as regras, códigos, princípios e eventos que contribuíram para a história da catalogação e dos catálogos. Assim, baseando-se em Moreno (2006) e Mey (1995), destacam-se:

a) Regras para catalogação - Anthony Panizzi

Podendo ser considerado o primeiro a sistematizar regras para catalogação, no período pré-catálogos, o bibliotecário inglês Anthony Panizzi redigiu as famosas 91 regras, publicadas em 1839 para o Museu Britânico. As suas regras tiveram grande influência na história da catalogação, inclusive, algumas delas ainda permanecem, como a valorização da página de rosto como fonte para identificar uma obra, por exemplo.

b) Código de Charles C. Jewett

Em 1850, Charles C. Jewett teve aceito seu código para o catálogo da Smithsonian Institution, nos Estados Unidos. Alguns preceitos de Jewett, em especial no tocante a cabeçalhos de responsabilidade e obras anônimas foram, por muito tempo, seguidos. Apesar de apresentar discordâncias e soluções diferentes, o código de Jewett baseou-se nas contribuições e regras de Panizzi, que teve seu mérito reconhecido no prefácio do seu código.

c) *Código de Cutter*

Em 1876, Charles Ami Cutter, considerado gênio por Ranganathan, publicou seu código, conhecido como *Rules for a dictionary catalog*, de fácil leitura, o qual constituiu-se um código muito completo. Segundo Mey (1995), a obra de Cutter é um exemplo de catalogação prática adequada ao usuário. Para ela, embora existam outros nomes e códigos no século XIX – época de proliferação de códigos de catalogação - Cutter marca esse período. Com a idealização do seu catálogo dicionário, observa-se uma mudança na lógica de organização, até então pautada, basicamente, pelo critério sistemático e que, a partir daí, passa a nortear-se, também, pela ordem alfabética (PINHO, 2006).

Apesar de ter criado, também, um esquema de classificação e uma tabela representativa de sobrenomes, para Mey (1995) a contribuição mais importante de Cutter está na catalogação.

d) *Código da American Library Association – ALA*

Em colaboração com a *Library Association* da Inglaterra, a ALA (*American Library Association*), publicou a

primeira edição do seu código - *Cataloging rules: author and title entries*, em 1908, utilizando regras de Panizzi, Cutter, Jewett e LC. O código da ALA estabelecia 369 regras, voltadas, especificamente, para entrada de autor e título na catalogação. De acordo com Corrêa (2008), este código foi editado três vezes para atender às solicitações dos bibliotecários, que não o aceitavam em sua totalidade, alegando, principalmente, o excesso de detalhamento de regras. A segunda edição, preliminar, do código da ALA, data do ano de 1941. A sua segunda edição definitiva foi publicada em 1949, e ainda sofreria muitas críticas, que apontavam para um acordo em termos internacionais e constituíram a base da Conferência de Paris.

e) Código da Vaticana

Na década de 1920, havia sido elaborado por bibliotecários norte-americanos, sob a direção de John Asteisson, norueguês de formação americana, o Código da Vaticana, que se tratava de esquemas para o catálogo de impressos, redigido especialmente para atender à reorganização da Biblioteca Apostólica da Vaticana. As normas da Vaticana basearam-se no código ALA de 1908, embora sua terceira edição tenha sido considerada superior a da segunda da ALA.

f) Princípios de Paris

Em 1961, a Declaração de Princípios da Catalogação, mais comumente conhecida por

Princípios de Paris, apresentou um resumo das regras desenvolvidas na Conferência Internacional sobre Princípios

de Catalogação, primeiro evento no sentido da normalização internacional. Essa conferência reuniu representantes de 53 países e de 12 organizações internacionais, com diferentes filosofias e códigos de catalogação dispostos a um acordo. O que se concretizou, a partir daí, foi a decisão sobre cabeçalhos para nomes pessoais e títulos uniformes. Os princípios decorrentes da conferência marcaram época por se tratarem de um primeiro esforço de colaboração em nível internacional e regem a catalogação até hoje, sendo usados, sempre que possível, como filosofia para códigos de várias nacionalidades.

g) Anglo-American Cataloging Rules – AACR

Em 1967, surge o *Anglo-American Cataloging Rules – AACR* (Código de Catalogação Anglo-Americano). Ao AACR seguiram-se uma 2ª edição, AACR2, publicada em 1978, também traduzida no Brasil (1983-1985), e que passou a ser adotada em quase todas as escolas de Biblioteconomia brasileiras, praticamente extinguindo a diversidade de códigos no ensino (MEY, 1995). Ao longo da história, o código passou por muitas alterações e revisões, no entanto, a versão mais recente da segunda edição do AACR2 surgiu em 2002, e foi publicada no Brasil em 2004 com o título Código de Catalogação Anglo-Americano (CCAA). As revisões das regras no AACR2 englobam proposições de outro esquema de descrição: o ISBD.

h) ISBD

O esquema de descrição ISBD - *International Standard Bibliographic Description* – (Descrição Bibliográfica Internacional Normalizada), proposto por um notável

especialista em catalogação, Michael Gorman, foi publicado pela IFLA (*International Federation of Library Association and Institutions*) em 1971, e teve seu início na Reunião Internacional de Especialistas em Catalogação – RIEC, realizada em Copenhague (1969), que trouxe mudanças significativas para os códigos e as práticas da catalogação. Esse documento sistematizou a ordenação das informações bibliográficas, identificando elementos e utilizando uma seqüência de pontuações padronizadas. É, até hoje, instrumento de comunicação internacional de informação bibliográfica. Outros documentos foram surgindo, inclusive, para contemplar diferentes tipos de suportes, como a ISBD(G), para variados tipos de materiais (MORENO, 2006).

i) Declaração dos Princípios Internacionais de Catalogação

A publicação da Declaração dos Princípios Internacionais de Catalogação, ocorrida em 2009, considerada por substituir e ampliar os Princípios de Paris, determinou, como fundamento para a construção de códigos de catalogação, **a conveniência do usuário**, ou seja, se estabeleceu que o usuário deve ser o foco primordial da atividade de catalogação. Dentro desse contexto, conforme aponta Zafalon (2012), cabe mencionar a repercussão da contribuição da autora Elaine Svenonius junto à publicação da Declaração dos Princípios Internacionais de Catalogação.

Zafalon (2012) destaca que Elaine Svenonius, que ao longo dos anos identificou lacunas quanto às características inerentes a um sistema bibliográfico, propôs ajustes aos objetivos bibliográficos definidos pela IFLA. E, nesse sentido,

apresentou um novo princípio, o da **navegação**, que objetiva fornecer melhorias tanto ao usuário quanto aos códigos, já que, a partir do ponto de vista de Elaine, os quatro princípios já existentes (localizar, identificar, selecionar e adquirir) mostravam-se como insuficientes. Dessa forma, segundo Brascher e Café (2008, p.5, grifo nosso), tem-se:

- **localizar** entidades em arquivo ou base de dados como resultado de uma busca por meio de atributos e relacionamentos entre as entidades;
- **identificar** uma entidade, isto é, confirmar que a entidade descrita em um registro corresponde à entidade desejada ou distinguir entre duas ou mais entidades com características similares;
- **selecionar** uma entidade que é apropriada às necessidades dos usuários;
- **adquirir** ou obter acesso à entidade descrita;
- **navegar** numa base de dados, isto é, encontrar obras relacionadas a determinada obra por meio de generalização, associação, agregação; encontrar atributos relacionados por equivalência, associação e hierarquia.

Cabe acrescentar que a autora Elaine Svenonius exerceu e exerce grande influência no domínio da organização da informação.

Diante dos fundamentos teóricos mencionados, percebe-se que houve uma sucessão de acontecimentos que, por vezes, estiveram interligados, empenhando esforços no intuito de fortalecer e propor melhorias que foram fundamentais para a evolução do processo de catalogação e para consolidação do tratamento descrito da informação. Nesse contexto, é possível destacar, após análise dessas contribuições, a existência de três aspectos fundamentais que caracterizam o TDI e que merecem uma atenção especial, a saber: processo, produto e instrumentos. Esses três aspectos

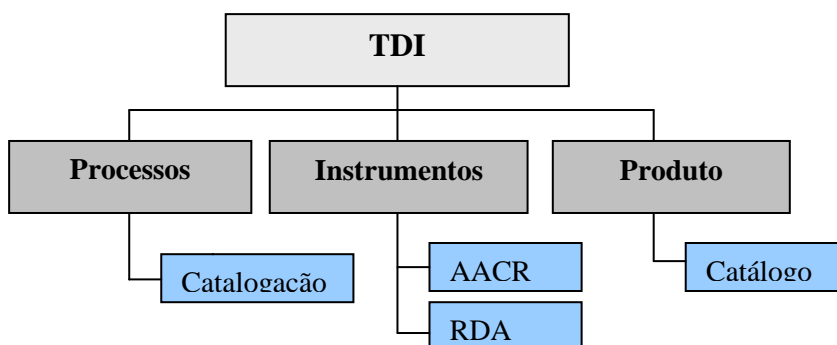
que configuram o TDI serão mais detalhadamente explicados na subseção a seguir.

3.1.1.2 Processos, produtos, instrumentos

Diante do exposto acerca do TDI, pode-se argumentar que ele se configura a partir de três aspectos que, juntos, compõem a sua essência; são eles: processo, produto e instrumentos. Assim, tem-se que o TDI desenvolve-se através de um processo (catalogação), que se vale de instrumentos (AACR, FRBR, RDA) para geração de um produto (catálogo).

Apresenta-se a figura 3 a seguir para evidenciar os três aspectos inerentes ao TDI.

Figura 3 – Aspectos inerentes ao TDI



Fonte: elaborado pela autora.

Desse modo, tem-se **a catalogação** que, enquanto processo, conforme abordado anteriormente, destina-se à representação sintética e codificada das características extrínsecas de um documento, ou seja, busca a

representação de um objeto informacional a partir de sua descrição física. Nesse aspecto, Corrêa (2008) afirma que a catalogação consiste em um mecanismo essencial para descrição e padronização de informações e que, para isso, “é construída a partir de regras que ofereçam o máximo de padronização e minimizem as interpretações individuais, procurando garantir a unicidade do item informacional representado e, ao mesmo tempo, sua universalidade” (CORRÊA, 2008, p. 20).

Entretanto, para que o processo de catalogação possa ser realizado utiliza-se de **instrumentos** de trabalho, que consistem em códigos e normas que norteiam o processo de catalogação. Tais instrumentos serão aqui destacados como: a) AACR; b) FRBR; c) RDA.

a) AACR - *Anglo-American Cataloging Rules* – (Código de Catalogação Anglo-Americano), consiste em um código ou conjunto de princípios que norteiam a catalogação, tendo como objetivo a normalização a âmbito internacional, subsidiando o tratamento da informação (ANZOLIN, 2007). Ao AACR, como visto em outra ocasião, seguem-se a sua segunda edição, o AACR2, e diversas revisões e alterações que ocorreram ao longo do tempo, sendo a mais recente datada de 2002, e publicada no Brasil em 2004.

O AACR2 é projetado para uso na construção de catálogos e outras listas em bibliotecas gerais. As suas normas abrangem a descrição e a provisão de pontos de acesso para os materiais de biblioteca, sendo dividida, assim, em duas partes: parte 1- lida com a prestação de informações que descrevem o item que está sendo catalogado; e parte 2 -

lida com a determinação e o estabelecimento das posições (pontos de acesso), a partir das quais a informação descritiva deve ser apresentada aos usuários do catálogo e com a realização de referências para aqueles títulos (ANGLO-AMERICAN CATALOGING RULES, 2006).

A esse respeito, Zafalon (2012) explica de forma mais detalhada que o AACR2 apresenta, por meio de orientações, regras e exemplos, a descrição do conteúdo (apresentada nas seções de parágrafos de um registro, que inclui o título, indicação de responsabilidade, edição, detalhes específicos do material, informação de publicação, descrição física, série, notas e números normalizados) e, também, a escolha, a elaboração e a atribuição de pontos de acesso a um documento (representando, por sua vez, nomes de pessoas, de lugares geográficos, instituições e eventos), o que permite que sejam criados os direcionamentos para a construção de catálogos bibliográficos.

c) O RDA - *Resource Description and Access* – (Descrição e Acesso ao Recurso), consiste em uma nova norma de catalogação, proposta pela IFLA, que pretende substituir a segunda edição do Código de Catalogação Anglo-Americano – AACR2. O RDA vai além dos códigos de catalogação anteriores ao prover orientações sobre como catalogar recursos digitais e auxiliar melhor os usuários para encontrar, identificar, selecionar e obter a informação desejada. Contribui, também, para o agrupamento de registros bibliográficos, visando mostrar relações entre obras e seus criadores. Essa importante e nova característica torna os usuários mais conscientes das diferentes edições,

traduções ou formatos físicos das obras – um significativo desenvolvimento. Dessa forma, o RDA se propõe a aumentar a eficiência na catalogação (JOINT STEERING COMMITTEE FOR DEVELOPMENT OF RDA, 2006).

A estrutura do RDA está construída a partir dos modelos conceituais do FRBR e FRAD⁷ para ajudar os usuários do catálogo a encontrarem mais facilmente a informação de que necessitam. Arellano (2009) enfatiza que as características e estrutura desse novo código exercerão influência na organização da informação, tendo um forte impacto nas tarefas de catalogação, assim como na forma e características do catálogo atual.

Após análise do processo de catalogação, que se vale de instrumentos para sua realização, tem-se como consequência o catálogo, visto como **produto** final da catalogação, utilizado para organização e recuperação de documentos, fornecendo auxílio ao usuário. O catálogo pode ser definido como um conjunto de registros bibliográficos de documentos de um acervo, organizados a partir de determinadas normas ou critérios.

Dessa forma, por meio dos catálogos, é dada a oportunidade aos usuários de encontrarem, identificarem, selecionarem e obterem os recursos bibliográficos necessários (ZAFALON, 2012). Essa abordagem é bem

⁷ O FRAD (*Functional Requirements for Authority Data*) complementa o FRBR na escolha, forma e controle de autores de qualquer natureza (autoridade). O FRAD é visto como um suporte para o RDA e relaciona os autores pessoais e/ou entidades autoras ao escopo do RDA, ou seja, contribui na identificação das relações que uma obra pode ter com seu criador (CORRÊA, 2008).

elucidada por Mey (1995, p. 9), ao informar que o catálogo “é um canal de comunicação estruturado, que veicula mensagens contidas nos itens, e sobre os itens, de um ou vários acervos, apresentando-as sob forma codificada e organizada, agrupada por semelhanças, aos usuários desse(s) acervo(s)”.

Diante do exposto, fica evidenciada a correlação entre processo, produto e instrumentos, tendo em vista que a catalogação, enquanto processo para descrição e representação de objetos informacionais, é construída a partir de instrumentos, que oferecem normas para padronização das representações e que, por fim, gera um produto, através do qual o usuário pode ter acesso ao conjunto de registros bibliográficos dos documentos de um acervo.

Após a compreensão do tratamento descritivo da informação, seus fundamentos teóricos, e os três aspectos que o norteiam (processo, produto e instrumentos), dedica-se a próxima subseção à compreensão do tratamento temático da informação.

3.1.2 Tratamento Temático da Informação

O Tratamento Temático da Informação (TTI) é voltado ao conteúdo informacional do documento (como na classificação, catalogação de assunto⁸, indexação e análise

⁸ Levando em conta a existência de duas dimensões da catalogação: a catalogação de forma (catalogação descritiva) e a catalogação de conteúdo (catalogação de assunto) (GUIMARÃES, 2009), se faz necessário esclarecer que a dimensão do tratamento temático da informação se encontra especificamente na catalogação de assunto,

documental), e tem como foco a identificação do objeto informacional com base em suas características intrínsecas, destinando-se à representação condensada do assunto com vistas a sua localização e acesso.

Acerca do TTI, Barité (1999, p. 124 apud GUIMARÃES, 2008) destaca que esse tratamento “tem por objeto os aspectos vinculados à análise, descrição e representação do conteúdo dos documentos, bem como suas inevitáveis interfaces com as teorias e sistemas de armazenamento e recuperação da informação”. Dessa maneira, pode-se argumentar que o tratamento temático tem como foco a representação, identificação e disponibilização do conteúdo informacional do documento, de modo que, a partir da representação gerada, o documento possa ser posteriormente localizado e acessado por um usuário.

Portanto, é possível compreender que o TTI cumpre um importante papel na interlocução entre os contextos de produção e uso da informação; tal fato é, também, observado por Guimarães (2008) ao enfatizar que o TTI ocupa (tanto na literatura quanto nas distintas práticas profissionais) um espaço nuclear, visto revelar a mediação entre a produção e o uso da informação, entre elas tecendo a mais sólida ponte: a que dá acesso ao conteúdo informacional.

Para melhor compreender os fatos que, ao longo da história, marcaram o desenvolvimento do TTI e que hoje fundamentam sua existência, apresenta-se, a seguir, uma

enquanto representação temática. Por outro lado, a dimensão do tratamento descritivo está na catalogação descritiva, que se refere à descrição bibliográfica.

breve fundamentação teórica acerca de alguns marcos conceituais.

3.1.2.1 Fundamentos teóricos

Percorrendo a literatura, Café e Sales (2010) trazem uma rica contribuição acerca dos pilares fundamentais que contribuíram para a construção da dimensão temática da organização da informação como conhecida hoje. Os marcos conceituais do TTI nos moldes atuais, manifestados notadamente na catalogação de assunto, na classificação, na indexação e na análise documental podem ser mencionados da seguinte maneira:

a) A Classificação Decimal de Dewey – CDD

Em momento histórico, em que a ideia de informação estava diretamente ligada ao livro (CAFÉ; SALES, 2010), o norteamericano Melvin Dewey, por meio de sua classificação decimal, no ano de 1876, ocupou-se em desenvolver uma lógica organizacional para acervos de bibliotecas baseada no assunto tratado nos livros, de modo a possibilitar a consulta (recuperação) deles de maneira uniforme. A principal contribuição de Dewey para a organização da informação reside na possibilidade de dar ordem aos assuntos (conhecimentos) sob uma perspectiva hierárquica, assuntos mais gerais e assuntos mais específicos.

b) Catalogação de assuntos de Cutter

O bibliotecário norte americano Charles Ami Cutter trouxe contribuições tanto para o tratamento descritivo (como

abordado na subseção 3.1.1.1), quanto para o tratamento temático da informação. Sua grande contribuição para o TTI encontra-se, também, na obra *Rules for a Dictionary Catalog* (1876), que inclui a abordagem da Catalogação de assunto. A catalogação de assunto enquanto abordagem teórica reflete uma tradição pragmática anglo-saxônica, centrada na geração de produtos advindos do tratamento temático da informação em bibliotecas. Ressalta-se que o catálogo dicionário de Cutter trouxe para o universo da catalogação uma nova roupagem, e se destacou por organizar os cabeçalhos em ordem alfabética como as palavras em um dicionário, fato que explica a origem do seu nome.

c) Classificação Decimal Universal – CDU

Em se tratando de organizar assuntos informacionais por meio de notações decimais, merece destaque o trabalho realizado pelos belgas Paul Otlet e Henri La Fontaine, criadores da Classificação Decimal Universal (CDU), em 1905. Lançando mão de codificações contidas de elementos semióticos, Otlet e La Fontaine concederam ao universo da organização temática, por meio da CDU, as primeiras formas de se efetuar uma organização (classificação) de assuntos com base na análise e síntese, cabendo a eles o pioneirismo da organização analítico-sintética, que surge conscientemente apenas na década de 1930 com Ranganathan.

d) A Indexação de Kaiser

O bibliotecário alemão Julius Otto Kaiser tem seu nome gravado nos escritos oriundos da classificação e indexação devido à publicação da obra *Systematic Indexing*, de 1911. Com a referida obra, em que o autor apresenta uma

maneira sistemática para a realização do processo de indexação de assuntos de documentos, o universo do tratamento temático da informação, até então familiarizado com a classificação e com a catalogação alfabética de assuntos, direciona também a atenção para a indexação alfabética, prática esta que foi revigorada com o sistema de Kaiser.

e) A classificação de Ranganathan

Entre 1933 e 1960, o bibliotecário e matemático indiano Shiyali Ramamrita Ranganathan, por meio da criação e do desenvolvimento da *Colon Classification* (CC), transcende o papel das classificações lineares e rigorosamente hierárquicas existentes até então, configurando um novo modo de classificar assuntos.

Algo semelhante, porém nem tão repercutido ocorreu, como já mostrado anteriormente, com Cutter – que criou um novo modo de formar cabeçalhos – e com Kaiser – que criou um novo modo de indexar assuntos. Portanto, a transformação ocorrida no universo das classificações, diante da publicação da *Colon Classification* é o efeito não do surgimento de uma nova classificação, mas da emersão de um novo método.

f) A teoria do conceito de Dahlberg

Não se pode tratar da dimensão temática da organização da informação sem mencionar a teoria do conceito de Dahlberg, cujas contribuições constituem um dos pilares fundamentais sobre os quais estão alicerçadas tanto a organização da informação quanto do conhecimento.

A autora alemã Ingetraut Dahlberg, na segunda metade do século XX, desenvolveu o estudo sobre conceitos na Ciência da Informação, preocupada com a compreensão da importância dos conceitos na representação do conhecimento. Sua teoria do conceito é descrita, também, como aporte teórico do tratamento temático para sistemas de informação (NONATO, 2009). Assim, a ideia de conceito está ligada às características dos objetos e, dessa forma, o conceito é considerado como uma série de enunciados (características) verdadeiros sobre um objeto, reunidos sinteticamente. Dahlberg (1978, p. 101) esclarece que:

todo enunciado sobre objetos contém um elemento do respectivo conceito, que se identifica como característica do conceito. Características idênticas evidenciam relações entre conceitos. A intensão de um conceito é a soma total de características e a extensão do conceito é a soma total de conceitos mais específicos. A categorização formal dos conceitos — objetos, fenômenos, processos, propriedades, relações — tem importância na formação de sistemas e na combinação dos mesmos.

Sem dúvidas, com o auxílio da teoria do conceito, ou seja, a partir da compreensão dos conceitos, suas características, categorias que se inserem, combinações e relações entre os mesmos, as atividades e sistemas voltados para representação e organização, seja da informação ou do conhecimento, podem ser elaborados de maneira mais objetiva.

g) O Classification Research Group (CRG).

Notória importância tem o trabalho realizado pelos pesquisadores do *Classification Research Group* (CRG), primeiro grande grupo de pesquisa no campo das

classificações, formado em Londres, em 1952, pelos professores A.J. Wells e B.C. Vickery. A equipe era constituída por professores, bibliotecários e documentalistas. Houve aqui um avanço das categorias do CRG com relação às categorias de Ranganathan, refletindo a mudança no universo da informação a ser organizada, visto que o foco dos estudos do CRG eram os assuntos especializados.

h) International Society for Knowledge Organization (ISKO).

Dando continuidade ao contexto temático da organização da informação, há que se destacar a *International Society for Knowledge Organization (ISKO)*, criada em Frankfurt, em julho de 1989, por Ingetraut Dahlberg, a partir dos trabalhos de mais de duas décadas da *Society for Classification* (GUIMARÃES, 2008). Sua base teórica reside, especialmente, nos princípios da classificação e nas pesquisas sobre tesouros, com particular destaque para os trabalhos de Ranganathan e do CRG.

A ISKO foi criada com base em duas vertentes da Biblioteconomia internacional, a primeira ligada ao estudo e desenvolvimento de tesouros como instrumentos alfabéticos de tratamento temático e, a segunda, vinculada à aplicação das teorias facetadas à elaboração de tesouros.

i) Análise documental.

No final da década de 1960, na França, J.-C. Gardin e M. Coyaude, preocupados com o desenvolvimento teórico-metodológico dos processos inerentes ao tratamento temático

da informação, dão início a uma visão mais abrangente, chamada *Analyse Documentaire* (análise documental).

Consistindo em uma operação intelectual de representação dos conteúdos documentais voltada à recuperação, a análise documental ancora-se em uma matriz de influência linguística. Os produtos da análise documental são os resumos, resultado do processo de condensação, e os índices, resultado do processo de indexação. A análise documental, de acordo com Café e Sales (2010), pode ser entendida como sendo a decomposição das partes de um todo, para maior compreensão do conteúdo informacional, para fins de representação e, conseqüentemente, de recuperação.

Observa-se, assim, que as habilidades que hoje se tem na OI, em específico, no âmbito do tratamento temático, são frutos de uma longa história tecida por profissionais e pesquisadores preocupados em tornar a informação acessível por meio de sua representação (CAFÉ; SALES, 2010).

Dessa forma, essa trajetória, além de revelar uma sucessão de descobertas, viabilizou, principalmente, a construção de correntes teóricas que, hoje, permeiam as discussões acerca do TTI.

3.1.2.2 Correntes teóricas

O desenvolvimento do TTI ao longo da história foi, pouco a pouco, sendo permeado por distintas concepções que, segundo Guimarães (2008, 2009) propiciaram o delineamento de três correntes teóricas, a saber: **catalogação**

de assunto (*subject cataloguing*), de orientação predominantemente norte-americana; **indexação** (*indexing*), de orientação predominantemente inglesa; e **análise documental** (*analyse documentaire*), de orientação predominantemente francesa. Essas correntes formam a base teórica e prática do tratamento temático da informação, resultando, assim, em uma base substancial da organização da informação.

A primeira corrente, a **catalogação de assunto**, é voltada diretamente para a atividade profissional em bibliotecas e sob forte influência da Escola de Chicago. A catalogação de assunto, de matriz norte-americana, tem seus princípios fundados na catalogação alfabética de Cutter e na tradição de cabeçalhos de assunto da *Library of Congress* (GUIMARÃES, 2008).

De acordo com Lancaster (2004, p.20), a catalogação de assunto “refere-se comumente à atribuição de cabeçalhos de assuntos para representar o conteúdo total de itens bibliográficos inteiros (livros, relatórios, periódicos etc.) no catálogo das bibliotecas”. Na mesma linha de pensamento, Silva e Fujita (2004) consideram a catalogação de assunto um método para expressar o conteúdo informacional do documento, usando, para isso, certo número de descritores. Dessa forma caracteriza-se, basicamente, como atribuição de cabeçalhos de assunto para representar o conteúdo total dos documentos em catálogos de bibliotecas.

A **indexação**, por sua vez, consiste na segunda corrente teórica do TTI. Essa concepção, de matriz inglesa, abrange, como colocado por Guimarães (2008), não apenas a

realidade bibliotecária tradicional, mas, inclusive, os centros de documentação especializados e o universo editorial, no qual os índices, enquanto produtos do TTI, decorrem da utilização de linguagens de indexação, notadamente os tesouros. Observa-se, assim, uma preocupação de natureza mais teórica acerca da construção de tais linguagens, em muito influenciada pelos trabalhos do *Classification Research Group*. Neste sentido, Café e Sales (2010) apontam que a indexação, enquanto linha de abordagem, tem como foco o desenvolvimento de instrumentos, ou linguagens de indexação, como os tesouros, que podem facilitar a pesquisa a documentos.

A **análise documental**, terceira corrente teórica do TTI, consiste em uma operação intelectual de análise e decomposição das partes de um todo, visando à melhor compreensão e representação dos conteúdos documentais e, conseqüentemente, possibilitando a recuperação. Como já mencionado em outra ocasião, essa linha de abordagem surgiu no final da década de 1960, na França, quando J.-C. Gardin e M. Coyaud, preocupados com o desenvolvimento teórico-metodológico dos processos inerentes ao tratamento temático da informação, deram início a uma visão mais abrangente, chamada *Analyse Documentaire* (análise documental). Dessa forma, a análise documental, como enfatiza Pinho (2010, p.37), “centra-se na explicação do próprio processo do tratamento temático que são os procedimentos que permitem a identificação e seleção de conceitos, gerando representação e produtos documentais”.

Assim, a partir da análise das três correntes teóricas do TTI, é possível destacar, de forma resumida, que elas apresentam efetiva complementaridade (inclusive histórica) e, em virtude de seus distintos objetos, revelam denominações distintas para fenômenos semelhantes (GUIMARÃES, 2009). Seguindo essa linha de pensamento, Café e Sales (2010) salientam que, embora essas três correntes se diferenciem quanto à ênfase dada aos seus fazeres, elas constroem juntas um arcabouço conceitual que, conjuntamente às ideias classificacionistas, formam a base do tratamento temático da informação.

Fundamentando-se na opinião dos autores citados anteriormente, nota-se que, aliada às correntes teóricas, a classificação também tem seu espaço no âmbito do tratamento temático da informação e, por isso, não pode deixar de ser citada. Esse pensamento é elucidado por Café e Sales (2010, p. 126, grifo nosso), ao destacar que:

o conjunto formado pelos estudos de catalogação de assuntos, **classificação**, indexação e análise documental forma a base teórica e prática do tratamento temático da informação, conseqüentemente, uma base substancial da organização da informação.

Nesse contexto, a classificação, enquanto atividade, exerce o papel fundamental de organizar e representar informação (SOUZA, 2000), para isso gera notações através de números ou símbolos que possibilitam uma representação e ordenação de documentos, facilitando sua disposição de maneira estruturada e auxiliando na sua recuperação.

Conceituando o termo “classificação”, Tristão, Fachin e Alarcon (2004) destacam que significa a ação e efeito de

classificar que, por sua vez, significa ordenar ou dispor em classes. Classificação significa, ainda, associar objetos idênticos e separar objetos diferentes. Assim a classificação, utilizada com frequência em bibliotecas e sistemas de informação, pode ser vista como o ato de agrupar documentos segundo suas características comuns, diferenciando-os dos demais, organizando-os em grupos ou classes e visando propiciar formas dinâmicas e diferenciadas de acesso a esses documentos.

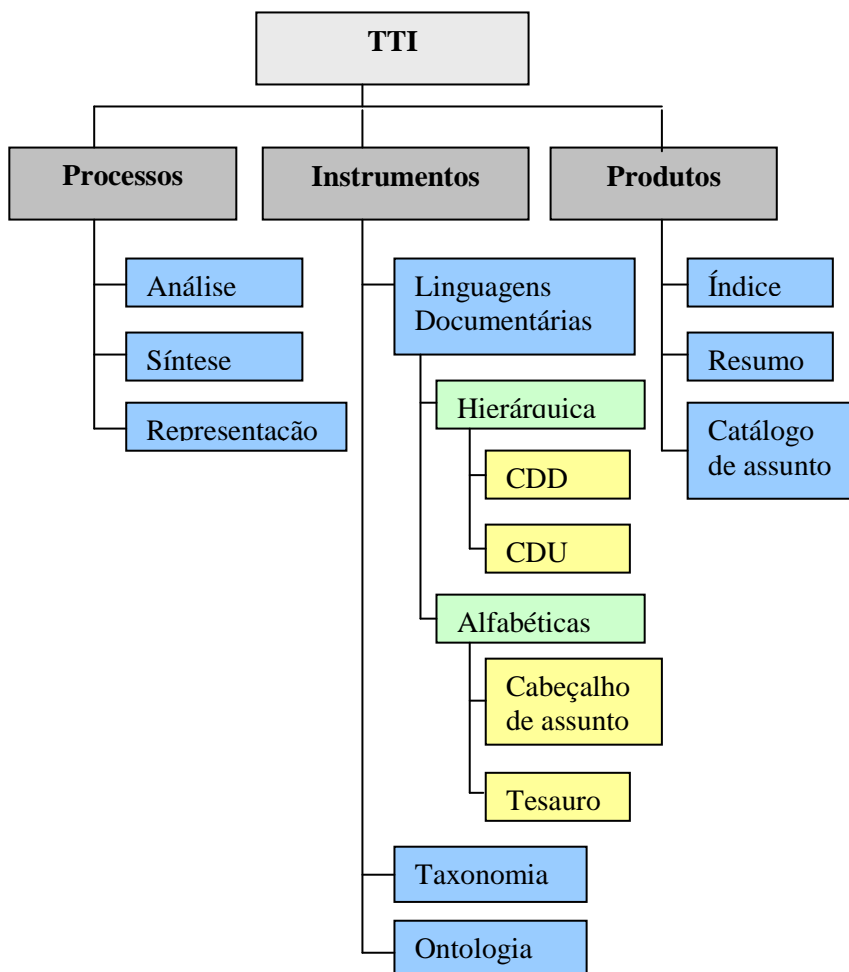
Dessa maneira, após análise das iniciativas que formam a base do TTI, mencionadas ao longo dessa subseção, torna-se importante destacar os três aspectos sob os quais se configuram os procedimentos inerentes ao TTI, a saber: processos, produtos e instrumentos.

3.1.2.3 Processos, Produtos e Instrumentos

Os três aspectos que configuram o TTI são descritos como: processos, produtos e instrumentos. Através desses três aspectos pode-se compreender, de forma prática, como se estruturam os **processos** que compõem o TTI; os **produtos** resultantes desses processos, que consistem em representações temáticas; e os **instrumentos** que fornecem aos processos as diretrizes necessárias para a sua realização.

A figura 4 evidencia, de maneira mais clara, os três aspectos inerentes ao TTI.

Figura 4 – Aspectos inerentes ao TTI



Fonte: elaborado pela autora.

Assim, são três os **processos** responsáveis pela realização do TTI; equivale dizer que as ações do TTI centram-se, de modo mais amplo, nas questões atinentes a: análise, síntese, e representação. A análise objetiva a

identificação de conteúdos informacionais, isto é, ocorre a partir da compreensão e da interpretação do conteúdo informativo de um documento (NAVES, 2001); a síntese, por seu turno, tem como função condensar o conteúdo do documento original e, ainda, segundo Galvão (1998), visa à construção de um enunciado, relativo à informação contida no documento e que será, posteriormente, submetido à representação temática; já a representação consiste na descrição das características de objetos informacionais ou documentos, de maneira que a representação gerada seja capaz de refletir a essência do documento representado.

Por outro lado, torna-se importante abordar os instrumentos inerentes ao TTI, pois para dar suporte ao desenvolvimento dos processos, anteriormente mencionados, o TTI se vale de **instrumentos**, que aqui estão estruturados, conforme visto anteriormente na figura 4, em: a) Linguagens documentárias; b) Taxonomias; c) Ontologias.

a) Linguagens documentárias.

As linguagens documentárias são um instrumento por meio do qual se realiza a mediação entre sistemas e usuários e designa, de modo amplo e completo, a linguagem especialmente construída para organizar e facilitar o acesso e a transferência da informação (LARA, 2004). Sendo assim, funcionam como linguagens de comunicação entre a informação documentária e o usuário que a demanda. Nesse contexto, as linguagens documentárias compreendem, genericamente, os sistemas de classificação, os cabeçalhos de assunto e os tesauros (MOREIRA; MOURA, 2006). Para melhor compreender, dividem-se, aqui, as linguagens

documentárias em dois grupos: linguagens *hierárquicas* e linguagens *alfabéticas*.

Como linguagens *hierárquicas*, podem-se citar os sistemas de classificação CDD (Classificação Decimal de Dewey) e CDU (Classificação Decimal Universal), que são sistemas artificiais de signos normalizados, que permitem representação mais fácil e efetiva do conteúdo documental, com o objetivo de recuperar a informação que o usuário deseja (TRISTÃO; FACHIN; ALARCON, 2004). Essas linguagens contribuem na construção de uma representação da informação no formato de números ou símbolos que, geralmente, descrevem o conteúdo do documento de forma mais abrangente.

Do outro lado, como linguagens documentárias *alfabéticas*, têm-se os cabeçalhos de assunto e os tesauros. Os cabeçalhos de assunto consistem em vocabulário controlado que padroniza e uniformiza a descrição sintética dos assuntos da obra no catálogo (MARTINHO; FUJITA, 2010). Assim, os cabeçalhos de assunto são elaborados a partir de palavras, geralmente extraídas do próprio documento, para descrever seus assuntos e funcionam, segundo Silva e Fujita (2004), como enunciados de assuntos, formados a partir da composição ordenada dessas palavras.

Os tesauros, por sua vez, consistem também, segundo sua estrutura, em vocabulário controlado, porém um vocabulário dinâmico de termos relacionados semântica e genericamente, cobrindo um domínio específico do conhecimento. O tesouro é, segundo sua função, um dispositivo de controle terminológico usado na tradução da

linguagem natural dos documentos, dos indexadores ou dos usuários, numa linguagem do sistema (linguagem de documentação, linguagem de informação) mais restrita (UNESCO, 1971).

b) Taxonomias

As taxonomias são instrumentos que organizam logicamente os conteúdos informacionais. Essa organização, segundo Tristão, Fachin e Alarcon (2004), remete para um procedimento classificatório e permite agrupamento categorizado, isto é, a partir de um assunto formam-se categorias que se dividem em classes e subclasses hierarquicamente, formando uma lista de categorias de assunto estruturada. Os autores observam, ainda, que as taxonomias adquirem importância como instrumentos que auxiliam na gestão do conhecimento e na organização e recuperação da informação, assegurando, no contexto de uma organização, que todas as aplicações da instituição utilizem a mesma linguagem para organizar, armazenar e apresentar a informação. Assim, como destaca Vital (2010), as taxonomias estão voltadas para a organização da informação em ambientes específicos visando à recuperação.

No domínio do TTI, as taxonomias são instrumentos que organizam logicamente os conteúdos informacionais, permitindo agrupamento categorizado e formando uma lista de categorias de assunto estruturada.

c) Ontologias

As ontologias, assim como as taxonomias e, também, os tesouros, tratam da representação e estruturação de

conceitos e seus relacionamentos em um domínio do conhecimento. A respeito das ontologias, Brascher e Carlan (2010, p. 160, grifo do autor) argumentam que elas “definem conceitos e relações de alguma área do conhecimento, de forma compartilhada e consensual e promovem e facilitam a interoperabilidade entre sistemas de informação, em um processo “inteligente” dos agentes (computadores)”. Ainda para Gruber (1993), uma ontologia é uma especificação formal e explícita de uma conceitualização compartilhada.

Assim, as ontologias estão focadas, também, no compartilhamento de informações e na interoperabilidade entre sistemas; porém, no âmbito do TTI, baseando-se em Silva (2008), entende-se que as ontologias servem como instrumentos para representação do conteúdo dos documentos, pois são constituídas por um vocabulário e um conjunto de definições consensuais em um domínio, que pode se apresentar em uma linguagem lógica.

Por fim, tendo analisado os instrumentos utilizados pelo TTI, vale mencionar que os três processos anteriormente citados contribuem para geração de **produtos** que têm como função facilitar a recuperação de informações relativas ao documento que foi representado. Tais produtos consistem em: a) Índice, b) Resumo, c) Catálogo de assunto.

a) Índice

O índice pode ser compreendido em dois sentidos: no sentido corrente e no amplo. No primeiro sentido, consiste em uma listagem alfabética ou sistemática de tópicos que indicam a existência e localização de cada um deles num documento ou em uma coleção de documentos. No segundo sentido, o

índice consiste em um conjunto ordenado de códigos representativos de assuntos, tópicos ou conceitos, os quais podem servir como critérios de busca, relacionando com alguma chave de acesso que permita localizar os documentos – ou suas partes ou representações – relativos a cada assunto, segundo Robredo (1994).

a) Resumo

Já o resumo, segundo produto do TTI, consiste em um produto decorrente da sintetização do conteúdo do documento. Para Lancaster (2004), o resumo é uma representação sucinta, porém exata do conteúdo de um documento, permitindo o acesso aos itens, ou complementando os pontos de acesso proporcionados pelos termos de indexação.

A norma atualmente adotada no Brasil para elaboração de resumo é a NBR 6028 (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 2003). Como bem colocou Café e Sales (2010), a norma determina três tipos de resumos:

1- Resumo indicativo: indica apenas os pontos principais do documento, não apresentando dados qualitativos, quantitativos, etc. De modo geral, não dispensa a consulta ao original.

2- Resumo informativo: informa ao leitor finalidades, metodologia, resultados e conclusões do documento, de tal forma que possa, inclusive, ser dispensada a consulta ao original.

3- Resumo crítico: resumo redigido por especialistas, com análise crítica de um documento. Também chamado de

resenha, quando publicado. Quando analisa uma determinada edição entre várias, denomina-se *recensão*.

Em relação às finalidades do resumo, Lancaster (2004) menciona que a mais importante está no fato de o resumo facilitar a seleção, ajudando o leitor a decidir, através da sua leitura, se determinado item apresenta a possibilidade de satisfazer seu interesse, poupando, dessa forma, o tempo do leitor. Assim, “em alguns casos, também, um bom resumo informativo pode realmente substituir a leitura de um item que seja de interesse para o usuário” (LANCASTER, 2004, p. 103). Outro aspecto em que os resumos são utilizados é no esclarecimento do conteúdo de documentos escritos em línguas que o leitor desconheça.

b) Catálogo de assunto

O catálogo de assunto é um produto organizado, como trata Silva e Fujita (2004, p. 142), “mediante determinação de cabeçalhos de assunto que funcionam como enunciados de assuntos formados a partir da composição ordenada de palavras”. Sua função é possibilitar ao usuário identificar documentos sobre uma determinada temática ou sobre temática relacionada ao que deseja. Nesse contexto, segundo Martinho e Fujita (2010), pode-se medir a eficácia do catálogo de assunto a partir da sua capacidade de conduzir os usuários às informações de que necessitam.

Diante do exposto acerca do tratamento da informação, tanto descritivo quanto temático, compreende-se que as questões que envolvem esses dois universos resultam, de forma geral, na produção de representações documentais, possibilitando a organização da informação e influenciando,

diretamente, na facilitação do acesso e recuperação de itens informacionais.

Nessa perspectiva, é importante considerar o tratamento da informação, enquanto aspecto essencial da OI, como uma atividade crucial dentro do contexto dos sistemas de recuperação de informação, tendo em vista que a organização da informação e a recuperação da informação possuem uma estreita ligação e podem trabalhar em parceria para gerar resultados satisfatórios. Assim, quando se trata de organizar informação para permitir a recuperação, quanto maior for a atenção voltada à organização da informação, mais eficaz poderá se tornar a recuperação.

Assim, considerando que os procedimentos de tratamento, enunciados até o momento, tiveram como foco o ambiente tradicional, torna-se pertinente abordar o procedimento a ser realizado quando se refere à organização da informação em ambiente digital.

3.2 Organização da informação em ambiente digital

O intensivo uso de tecnologias de informação e comunicação (TICs) e seus impactos globais contribuem para caracterizar a sociedade atual como uma sociedade da informação, que vive a era digital, em que o acesso à informação não enfrenta mais barreiras físicas. Esse novo contexto trouxe significativas mudanças, como destaca D'Andréa (2006, p. 39):

A emergência da chamada sociedade da informação acelerou um processo de mudanças socioculturais

caracterizado, entre outros fatores, pela rápida disseminação de informações e facilidade de comunicação em nível global, principalmente por meio das chamadas tecnologias da informação e comunicação (TICs). Com o advento dessas ferramentas, a publicação e o acesso à informações de diferentes naturezas passam a ocorrer em escala nunca antes vista.

Nesse contexto, está inserida a internet, peça chave que trouxe grande contribuição para consolidação da sociedade da informação, possibilitando o aumento da produção e disponibilização de informações em ambiente digital. A Internet pode ser definida como a rede das redes ou como uma rede interativa de redes, agregando serviços e recursos interativos, sem que possa ser fisicamente visualizada ou localizada em termos de espaço e tempo (SOUZA, 2000). Assim, a internet abriu novos horizontes para a disseminação da informação, impulsionou ainda mais a troca de informações e a diminuição de distâncias, permitindo a comunicação entre pessoas localizadas em qualquer parte do mundo.

Como consequência, os sistemas que lidam com informação também se adequaram a essa nova realidade, isto é, ao novo ambiente digital, que possibilitou desenvolvimentos significativos na produção, no acesso, disseminação e recuperação de informação. No entanto, essa nova realidade, apesar de trazer grandes benefícios, trouxe, também, grandes desafios.

Acerca disso Feitosa (2006) comenta que o crescimento da qualidade de informação eletrônica, disponível por meio da Internet, é uma realidade que pode ser percebida

em todos os campos do conhecimento humano, mas, por outro lado, devido ao crescente volume de informações, tornam-se maiores os problemas de acesso e de recuperação de qualquer tipo de informação na rede. Wives (1999) destaca o assunto ao comentar que o aumento informacional, assim como as dificuldades por ele acarretadas, crescem com o aparecimento da Internet, ampliando a sobrecarga da informação, onde o crescente volume de dados não favorece a recuperação de uma informação útil.

Diante desses desafios, é prudente observar a necessidade de organização da informação disponível nos sistemas em ambiente digital, pois a organização dessas informações pode interferir diretamente na sua localização e recuperação.

Um primeiro ponto a ser observado, antes de adentrar na organização da informação, refere-se ao documento digital, ou seja, aos documentos que comportam a informação no ambiente digital. Esses documentos podem ser vistos como documentos em formato eletrônico, que podem ser acessados por meio de uma interface ou tecnologia específica (MONTEIRO, 2010). Os documentos digitais, assim como os impressos, merecem um tratamento especial para que sejam satisfatoriamente utilizados.

Em se tratando da organização da informação a ser feita no ambiente digital, em que se inserem os sistemas que lidam com a informação, aliados aos documentos em formato eletrônico, se faz necessário que cada objeto informacional seja representado adequadamente, assim como ocorre no ambiente tradicional. A representação dos objetos

informativos é um procedimento fundamental e que influencia na recuperação da informação, uma vez que representa e identifica as características do documento, fornecendo uma síntese do seu conteúdo e estabelecendo categorias para seu armazenamento e posterior localização.

Como visto na subseção anterior, a representação pode ser realizada através do tratamento físico e temático da informação. O tratamento físico, voltado aos elementos relativos à forma dos documentos, com base na análise do seu tipo e identificação das informações descritivas. E o tratamento temático, voltado ao conteúdo informativo, buscando representar de forma condensada o conteúdo dos documentos. Portanto, o que se pretende analisar aqui é como adequar os procedimentos de representação da informação, anteriormente enunciados, à organização da informação no ambiente digital.

Dessa forma, tendo em vista que a OI compreende a representação dos documentos de acordo com seus aspectos físicos e temáticos, no ambiente digital, essa representação pode ser realizada com metadados, que se referem, etimologicamente, a “dados sobre dados, apontados pela literatura científica com o objetivo de descrever e localizar documentos eletrônicos na Web, de forma a permitir a sua recuperação de forma mais adequada e precisa” (CASTRO; SANTOS, 2010, p. 304).

De uma forma simples, Monteiro (2010, p. 134) define metadados como “dados codificados e estruturados que descrevem as características de um recurso, seja este documento, produto ou serviço”. Rowley (2002) enquadra

como tipos de metadados os registros bibliográficos, que vêm sendo empregados no contexto mais especializado de dados que se referem a recursos digitais disponíveis numa rede.

Acerca dos metadados, Feitosa (2006) assinala que os elementos de um registro de metadados associam-se a propriedades como nome, domínio, tipo, tamanho, formato, responsabilidade e origem. O autor, utilizando-se da contribuição de Day (2000 apud FEITOSA, 2006), apresenta, ainda, alguns dos benefícios decorrentes do uso de metadados, como: a) metadados são estruturados e, desse modo, podem formar a base para o desenvolvimento de sistemas de busca mais avançados; b) metadados podem ser convertidos para outros formatos, de modo que possam interoperar com diferentes protocolos de busca e recuperação; c) em relação a documentos publicados na internet, pensando-se em recuperação automatizada, torna-se mais fácil a extração de conteúdo de uma base de dados de descrição de recursos, do que de arquivos textuais em formato HTML; d) metadados administrativos podem ser utilizados para facilitar o gerenciamento do sistema de informação, uma vez que ajudam a avaliar quando os recursos devem ser revistos ou removidos da base de dados.

O conjunto de dados essenciais, utilizados como padrão para tornar possível a descrição de recursos na web, são os chamados esquemas ou padrões de metadados. Entre os padrões de metadados existentes vale destacar o Dublin Core como uma das iniciativas mais populares. O Dublin Core é mencionado por Monteiro (2010) como o conjunto mínimo de metadados, requerido pelo *Open Archives Initiative* -

Protocol for Metadata Harvesting (OAI-PMH), planejado para facilitar a descrição de documentos digitais. De acordo com Feitosa (2006, p. 54):

O esquema Dublin Core caracteriza-se pela sua utilidade e flexibilidade na representação de dados. A semântica dos elementos foi projetada para ser claramente compreendida por um grande número de pessoas, sem a necessidade de treinamento, bem como para interagir com o maior número possível de tecnologias. Assim, a intenção da metodologia não é suplantiar outros recursos usados para descrever documentos, mas, ao contrário, complementa-los.

O padrão de metadados Dublin Core possui, como principais características, a simplicidade do uso, entendimento semântico universal, escopo internacional e extensibilidade, que permite a adaptação às necessidades adicionais de descrição (DUBLIN CORE METADATA INITIATIVE, 2004). Devido à proliferação dos documentos eletrônicos e, em face da impossibilidade de os profissionais especialistas em informação realizarem a catalogação de todos os documentos, o padrão Dublin Core, que foi concebido originalmente para a descrição de documentos na web, pensou em um conjunto de treze elementos para descrição dos documentos, passando, posteriormente, para quinze elementos: título, criação, assunto, descrição, editor, colaborador, data, tipo, formato, identificador, fonte, idioma, relação, cobertura e direitos (FEITOSA, 2006).

Dessa forma, semelhantemente aos tratamentos descritivo e temático realizados para representar um documento no ambiente tradicional, visando descrever as suas características extrínsecas e intrínsecas, também ocorre com a descrição de um documento no ambiente da web; no

entanto; tal descrição é realizada através dos elementos acima citados, ou seja, dos campos capazes de descrever as principais informações de um registro, visando sua representação.

Outro aspecto a ser citado no contexto do ambiente digital está relacionado aos instrumentos que podem auxiliar na organização da informação. Os instrumentos que podem servir como suporte à organização da informação no ambiente digital foram, também, abordados na subseção 3.1.2.3, são eles: **tesauros**, que consistem em vocabulários dinâmicos de termos relacionados semântica e genericamente, cobrindo um domínio específico do conhecimento; **listas de cabeçalhos de assunto**, que são listas de enunciados de assuntos formados a partir da composição ordenada de palavras; **taxonomias**, como instrumentos que organizam logicamente os conteúdos informacionais; **ontologias**, que definem conceitos e relações de alguma área do conhecimento, de forma compartilhada e consensual. Assim, tais instrumentos possibilitam que as informações representadas por metadados sejam submetidas a uma linguagem comum ou padronizada, a fim de que uma comunicação efetiva aconteça.

A utilização desses instrumentos permite, ainda, quando se trata de tesauros, taxonomias e ontologias, evidenciar as relações existentes entre os conceitos usados para representar os assuntos dos documentos. Essa evidenciação é possível, tendo em vista que esses instrumentos possibilitam uma representação sistemática e estruturada de conceitos e seus relacionamentos, o que pode facilitar na localização e recuperação de documentos. Dessa

forma, a adoção de instrumentos auxilia tanto na organização da informação, quando na recuperação da informação, atuando, também, como suporte à navegação no ambiente web.

Ao discorrer sobre o ambiente digital e o uso de metadados, cabe, aqui, com base em Monteiro (2010), fazer uma analogia ao ambiente tradicional, assim, pode-se entender que a principal diferença entre o ambiente tradicional e o digital está no acesso aos conteúdos, uma vez que na biblioteca tradicional a representação remete à localização física de um documento impresso disposto na estante; já a representação da informação no ambiente digital indica ligações (links) para acessar documentos digitais em um repositório digital.

Esse fato remete a necessidade da adoção de interfaces que auxiliem no processo de busca, pois nos sistemas de informação em ambiente digital o usuário é desprovido da possibilidade de localização da informação impressa que ocorre em ambientes físicos, como na biblioteca. Aí encontra-se o importante papel da interface gráfica, ao servir de intermédio para o acesso às coleções digitais destes sistemas. Sobre esse tema, tem-se a colaboração da área de Visualização da Informação, que pode ser considerada como um meio de potencializar o acesso e recuperação de informações.

4 VISUALIZAÇÃO DA INFORMAÇÃO

A VI contribui, por meio da construção de interfaces gráficas, para representação visual de conteúdo informacional e atua no intuito de facilitar o acesso às informações que foram previamente organizadas.

Nesse sentido, a VI pode ser definida como uma área da ciência que tem por objetivo o estudo das principais formas de representações gráficas para apresentação de informações, a fim de contribuir para o melhor entendimento delas, bem como ajudar a percepção do usuário na dedução de novos conhecimentos baseado no que está sendo apresentado (FREITAS et al., 2001).

Outra interessante definição é trazida pelo grupo de pesquisa em VI do instituto de pesquisa de *software* da Universidade da Califórnia, Irvine⁹, segundo o qual a VI está centrada no desenvolvimento e análise de métodos para a apresentação de informações abstratas de forma visual. Tendo em vista que a representação visual de informações permite que as pessoas se tornem mais facilmente consciente de fatos essenciais, desenvolvendo uma compreensão mais profunda sobre o que está sendo apresentado e permitindo, através de uma visualização interativa, manipular e explorar informações.

Resumidamente, a VI está fundamentada no fato de que a informação apresentada de maneira visual possibilita uma compreensão mais clara daquilo que está sendo

⁹ <http://www.isr.uci.edu/research.html>

transmitido, já que, como acrescenta Dias e Carvalho (2007, p.1):

na maioria dos casos o oferecimento de imagens, figuras, estruturas gráficas e quaisquer outros recursos gráficos, com a finalidade de apresentar uma informação, produz a compreensão da mensagem transmitida, pois esta se torna mais natural e exige menos esforço cognitivo.

Esse ponto de vista também é compartilhado por Zhang (2008) ao relatar que os seres humanos são atraídos por imagens e informação visual. Segundo Zhang, as imagens e gráficos podem ser cativantes e atraentes, especialmente, se bem elaborados, pois uma representação visual pode comunicar informações de maneira mais rápida e eficaz que qualquer outro método. Como exemplo básico, pode-se considerar a diferença entre uma descrição escrita do rosto de uma pessoa e uma fotografia dela, nesse caso a informação transmitida pela imagem poderá gerar uma compreensão mais clara e precisa acerca da pessoa à quem se refere do que a informação descritiva.

Estivalet (2000) também reafirma essa idéia ao ressaltar que se torna fácil para uma pessoa capturar o significado de uma imagem ao invés de vários dados isolados, pois uma representação visual proporciona uma forma simples e intuitiva de entender o significado das informações a serem transmitidas.

Nesse aspecto, é importante destacar que, para que seja possível a construção dessas representações visuais, a VI trabalha em parceria com as ferramentas computacionais, que se configuram como facilitadoras dos processos de visualização. A esse respeito, Mazza (2009) explica que,

embora a visualização não dependa necessariamente de computadores para cumprir seu papel, por ser um processo que acontece na mente, as ferramentas computacionais advindas do avanço da tecnologia tem o poder de potencializar o uso de visualizações, através de recursos modernos e interativos.

Ware (2004) e Card et al (1999 apud PRESSER; SILVA; SANTOS, 2010) pontuam, nesse sentido, que a VI procura reunir ambos, o poder da percepção visual humana e o poder de processamento do computador, para analisar e compreender um conjunto de informações. Assim, com base nessa colocação, pode-se entender que a VI, tendo como objetivo final a inclusão informacional dos usuários, conta, de maneira geral, com o auxílio da tecnologia e dos computadores para geração de representações gráficas, se apoiando no uso da capacidade visual humana que, como afirma Nascimento e Ferreira (2005), é o sentido que possui maior capacidade de captação de informações por unidade de tempo.

Como visto até o momento, a VI, enquanto atividade temática, está voltada para as questões que envolvem a otimização na representação e, conseqüentemente, no acesso e recuperação de informações. Desse modo, cabe aqui retomar a justificativa desta pesquisa, que define que a VI se situa como um meio de potencializar a apropriação de informação, por meio da construção de interfaces gráficas para sistemas de recuperação de informação. Assim, dirigindo-se a visualização, a partir da perspectiva da recuperação de informação, é possível analisar o quanto às

características visuais, perceptivas e cognitivas da VI podem trazer diversos benefícios à localização e recuperação da informação (tais benefícios serão abordados na seção 4.1 a seguir).

Buscando compreender de forma mais clara a influência que a visualização exerce na recuperação de informação, se faz relevante destacar a visão de Zhang (2008) ao elucidar que o dinamismo, a diversidade, a heterogeneidade e a complexidade de informações dispostas na internet e em outros sistemas de recuperação de informação têm representado um desafio para as técnicas e teorias tradicionais de recuperação de informação. Esses desafios levam a necessidade de adoção de uma maneira mais interativa, intuitiva e eficaz de recuperação de informação nos sistemas. Nessa perspectiva, encontra-se a VI que, por meio das interfaces gráficas, se configura como uma poderosa aliada em SRIs.

Assim, quando tais sistemas possuem suas informações devidamente organizadas, a utilização de interfaces gráficas pode vir a ser um potencializador na representação e, conseqüentemente, na recuperação, minimizando algumas das dificuldades enfrentadas pelos sistemas, já que o uso dessas interfaces facilita a interação com o usuário e permite, como afirma Mazza (2009), entender sistemas complexos, tomar decisões e descobrir informações importantes que, de outra maneira, poderiam permanecer ocultas.

Além do mais, a adoção de interfaces gráficas por SRIs tem se tornado cada vez mais acessível, pois o avanço

das tecnologias de informação, dos equipamentos de imagens em geral e dos computadores tem tornado possível a construção de sistemas complexos e ricos na qualidade da apresentação de suas informações, através de recursos gráficos e de interfaces cada vez mais interativas e de fácil utilização.

4.1 Benefícios trazidos pelo uso da Visualização da Informação

De maneira prática, a utilização de interfaces de VI traz consigo grandes benefícios aos sistemas que as adotam. Podendo ser considerada uma poderosa alternativa para representação, acesso e recuperação de informação, as interfaces gráficas se configuram como uma forte aliada para os sistemas, pois:

a) Possibilitam a compreensão de como está estruturada a informação

A interface gráfica tem a capacidade de oferecer uma visão geral dos assuntos tratados, tornando possível visualizar, na maioria dos casos, como estão estruturadas as informações e quais as relações estabelecidas entre elas no sistema. Desse modo, as estruturas de VI ampliam a compreensão do usuário acerca das informações transmitidas e, como afirma Robertson, Card, Mackinlay (1993), tentam exibir as relações estruturais e de contexto que seriam mais difíceis de detectar por solicitações individuais de recuperação.

b) Facilitam a interação com o usuário

As interfaces baseadas em visualizações oferecem ao usuário a possibilidade de interação com o sistema, ou seja, o usuário tem a possibilidade de se comunicar por meio de uma interface amigável enquanto procura informações de seu interesse. Em alguns casos a interação pode se dar no sentido de controlar a quantidade de informações a serem exibidas na tela, ou mudar a forma como essas informações são apresentadas visualmente. Os mecanismos de interação objetivam proporcionar ao usuário a manipulação das representações visuais, permitindo a exploração de diferentes cenários para um melhor entendimento das informações visualizadas.

c) Auxiliam em buscas efetivas

Nesse aspecto, a utilização de uma interface gráfica tem a capacidade de orientar o usuário para a informação desejada de forma simplificada, acarretando em buscas mais efetivas. Como se sabe, a formulação de uma busca ocorre, geralmente, por meio de palavras-chave em uma expressão de busca, o que se configura em uma tarefa difícil, pois, por vezes, o usuário não tem sua necessidade bem esclarecida e enfrenta barreiras no momento de traduzir essa necessidade para o sistema.

Dessa forma, como explica Baeza-Yates e Ribeiro Neto (1999), a busca por informação é um processo impreciso, pois quando os usuários se aproximam de um sistema de acesso à informação, muitas vezes, tem apenas um conhecimento difuso de como podem alcançar seus objetivos. Nesse cenário, a interface visual tem o poder de

auxiliar por meio de buscas exploratórias, oferecendo uma visão mais ampla dos assuntos tratados e possibilitando uma navegação amigável, onde o usuário tem a oportunidade de percorrer o sistema e selecionar informações que sejam, de fato, relevantes a sua necessidade.

Desse modo, se torna menor o esforço do usuário na tentativa de expressar uma necessidade apenas por meio de palavras-chave. No entanto, ainda assim, algumas visualizações apresentam, também, as palavras-chave como possibilidade de realizar buscas, nesses casos há um direcionamento claro que conduz o usuário pela interface até a localização da informação pesquisada.

d) Permitem o descobrimento de novas informações

A utilização de interfaces gráficas, além de permitir a compreensão de como está estruturada a informação, facilitar a interação com o usuário e auxiliar numa busca efetiva, se configura, ainda, como um meio poderoso de promover a descoberta de informações extras, cuja existência, até então, era desconhecida pelo usuário. Dessa forma, o usuário, ao navegar no sistema em busca de itens que satisfaçam sua necessidade, tem a possibilidade de ser surpreendido com novas informações que não eram seu objeto de busca, mas que podem vir a ser úteis. Assim, como menciona Zhang (2008), a VI, por meio de interfaces gráficas, oferece uma maneira única para revelar informações ocultas.

Diante do exposto, convém elucidar que os benefícios trazidos pelo emprego de interfaces de VI auxiliam desde a compreensão das informações apresentadas, potencializando o acesso e a localização, até a recuperação de informações

relevantes. Além disso, as interfaces gráficas, por serem de fácil uso, atendem aos mais variados públicos, sendo úteis tanto para usuários experientes como para os que não têm experiência com recursos visuais, pois é precisamente a característica visual o ponto chave que torna as interfaces acessíveis.

No entanto, para que as visualizações cumpram seu objetivo de maneira satisfatória, necessitam ser dotadas de dois importantes atributos, sob os quais as visualizações são analisadas para comprovar sua utilidade e bom desempenho. Tendo como base à contribuição de Nascimento e Ferreira (2005), destacam-se os dois atributos, a saber: **expressividade** e **efetividade**.

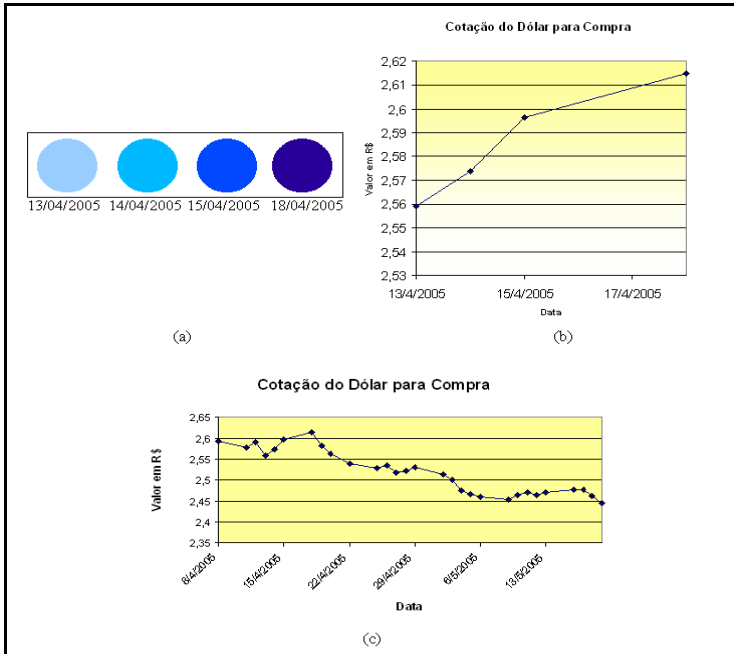
Uma visualização é considerada **expressiva** quando é capaz de expressar todas as informações de interesse do usuário e nenhum dado a mais ou a menos fora do contexto. Assim, quando a visualização atende ao usuário que a ela se destina, fornecendo informações relevantes que façam sentido e evitando informações desnecessárias, tem-se uma visualização expressiva.

Por outro lado, é considerada **efetiva** a visualização que facilita a compreensão dos dados apresentados pela estrutura, de forma que o usuário possa localizar a informação desejada sem erros de interpretação. Dessa maneira, tem-se a efetividade de uma visualização ao passo que se garante um fácil entendimento por parte do usuário.

Com o intuito de possibilitar uma melhor compreensão acerca da análise da expressividade e da efetividade em visualizações, apresenta-se a seguir, baseando-se em

Nascimento e Ferreira (2005), a figura 5, que é composta por três visualizações (a), (b) e (c), utilizadas para apresentar a variação da cotação do dólar para compra em reais.

Figura 5 - Visualizações da cotação do Dólar para compra



Fonte: Nascimento e Ferreira (2005, p. 1272)

A visualização (a) mostra a cotação no período de 13 à 18/04/2005, utilizando um mapeamento visual em que o valor do dólar foi associado à saturação da cor utilizada em uma sequência de círculos, nessa visualização para cada dia útil do período há um círculo correspondente, quanto mais escuro é o círculo, maior é a cotação do Dólar. Como se pode perceber, essa visualização, além de não ser expressiva, pois

oculta os dados necessários para o entendimento completo da informação transmitida, ou seja, os valores exatos da cotação; também não é efetiva, já que o esquema de cores pode causar erros na interpretação dos dados.

A visualização mostrada em (b), onde os dados são apresentados intuitivamente através um gráfico XY de linhas, tem sua expressividade afetada, pois, apesar de, aparentemente, apresentar os dados necessários para a compreensão da informação que está sendo transmitida, o gráfico dessa visualização carece de dados para situar o período analisado em um contexto maior. Fato que afeta, também, sua efetividade, pois essa visualização pode sugerir que houve uma forte tendência de aumento do dólar, uma vez que a linha de projeção da cotação é íngreme.

Por outro lado, na visualização (c), onde é demonstrada a situação do dólar entre 08/04/2005 e 13/05/2005, percebe-se que houve, na verdade, uma queda crescente do seu valor. Desse modo, pode-se afirmar que as duas primeiras visualizações não são, portanto, expressivas, por não expressarem os dados suficientes para compreensão, e nem efetivas, porque não conseguem transmitir para o observador uma informação correta sobre a tendência significativa da variação do dólar. Já a visualização (c) pode ser considerada expressiva e efetiva, quando comparada com as demais, pois oferece os dados necessários para uma compreensão correta do que está sendo transmitido.

Dessa forma, expressividade e efetividade são aspectos importantes que devem estar presentes em toda estrutura de VI, pois sem eles uma visualização pode não ser

capaz de transmitir dados relevantes, não trazendo, assim, informações novas além daquilo que já é conhecido. Ou, além disso, uma visualização pode ser, também, de difícil entendimento, podendo induzir o usuário a uma interpretação errônea, ou dificultando a localização da informação desejada.

Nesse contexto, se faz necessário um planejamento prévio, que envolve a escolha de técnicas adequadas que melhor se adaptem aos objetivos e ao contexto em que estão inseridos os sistemas e, também, aos usuários que pretende atender, pois, como bem afirma Estivalet (2000), é necessário entender o sistema e as necessidades dos usuários de forma a criar representações visuais efetivas, ou seja, é preciso compreender o propósito primeiro para depois usar a visualização que melhor se adequa.

4.2 Técnicas de Visualização da Informação

A VI congrega um conjunto de técnicas que tem por objetivo facilitar o entendimento de informações a partir de representações visuais (ESTIVALET, 2000). Essas técnicas de VI procuram representar graficamente informações, a partir de mecanismos interativos, de modo que o usuário possa compreender o conjunto de informações apresentadas.

A utilização das técnicas de VI para representação de informações permite ao usuário perceber, interpretar e compreender as informações apresentadas de maneira transparente e interativa, isso se dá pelo fato de que a adoção

dessas técnicas de visualização, como destacam Nascimento e Ferreira (2011, p. 16), “tem um forte apelo quando comparada com outras formas de transmitir e/ou analisar informações”.

Nesse aspecto, essas técnicas podem se apresentar como um poderoso meio de facilitar o entendimento e o acesso a informações em SRIs, uma vez que a disposição de informações por meio de recursos gráficos, como abordado anteriormente, exige menor esforço na compreensão, e aumenta as chances de localização de informações precisas.

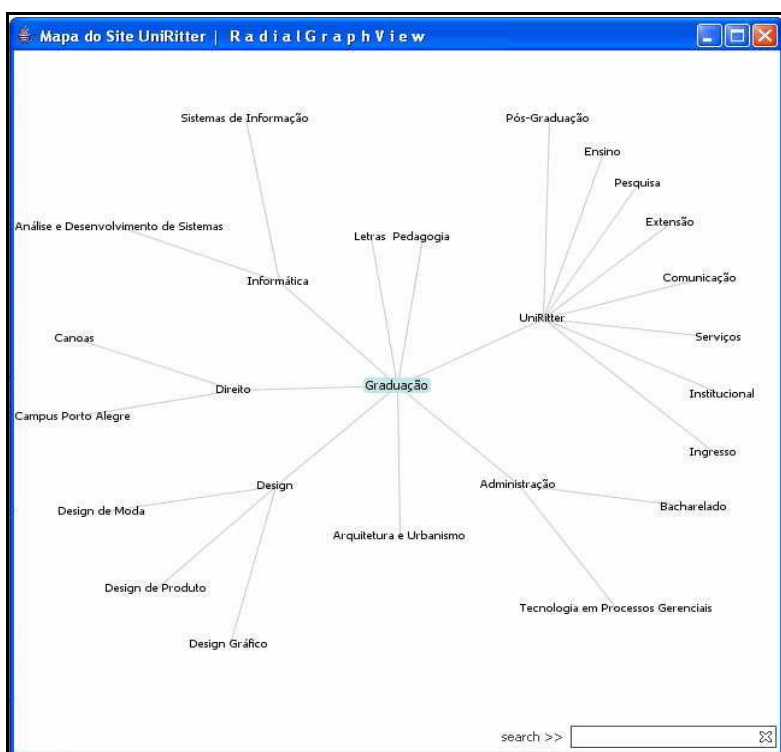
Desse modo, buscou-se, aqui, abordar algumas técnicas, entre as diversas existentes, que melhor atendam ao propósito acima mencionado, ou seja, técnicas voltadas para representação de informações que tenham aplicações práticas na construção de interfaces para sistemas que lidam com a recuperação da informação. Assim, será feita a seguir uma explanação dessas técnicas e, em seguida, de alguns modelos de interface visual que podem ser empregados em sistemas de informação.

a) Desenho de Grafos

A técnica desenho de grafos é formada, como o próprio nome evidencia, por grafos, que são modelos matemáticos formados por estruturas simples que consistem de um conjunto de vértices e um conjunto de arestas, os vértices geralmente representam objetos concretos ou abstratos e as arestas indicam relações entre esses objetos (NASCIMENTO; FERREIRA, 2011). Os grafos podem ser

A figura 7 também consiste em uma representação que utiliza a técnica desenho de grafos, nesse caso a técnica foi aplicada para estruturar o conteúdo de um site, onde os vértices representam os cursos oferecidos por uma instituição de ensino e as arestas indicam a relação estabelecida entre esses vértices (cursos).

Figura 7 - Desenho de Grafos

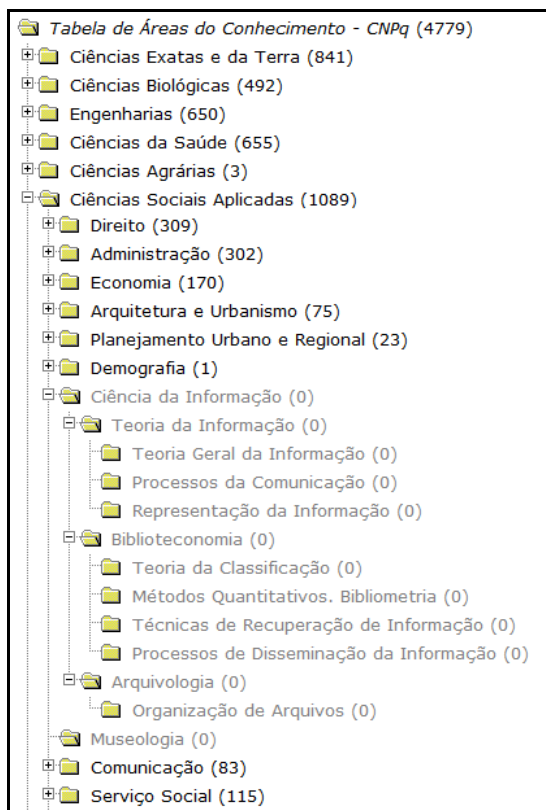


Fonte: Souza (2007)

b) Desenho de Árvores

O desenho de árvores consiste em uma particularização do desenho de grafos. Dentre as classes estudadas em desenho de grafos, o desenho de árvores está entre as primeiras, em função da sua popularidade e da simplicidade de sua estrutura (NASCIMENTO; FERREIRA, 2005).

As técnicas para desenhar árvores consideram dois casos distintos: o desenho de árvores com raiz (do inglês *rooted trees*), que representam hierarquias, tais como diagramas organizacionais, árvores de diretórios e taxonomias; e o desenho de árvores livres (do inglês *free trees*), onde as árvores não possuem raízes, mas, da mesma forma, são capazes de representar hierarquias por meio de um desenho radial. Em ambos os casos, é desejável a obtenção de desenhos planares em linhas poligonais (NASCIMENTO; FERREIRA, 2005). Como exemplo de uma estrutura de desenho de árvores tem-se a figura a seguir, onde estão estruturadas as áreas do conhecimento do CNPq.

Figura 8 – Desenho de árvores das áreas de conhecimento do CNPq

Fonte: Vieira (2011, p. 15)

c) Browser Hiperbólico

O browser hiperbólico consiste em outra importante técnica, muito utilizada por sistemas na web. Essa técnica apresenta hierarquias por meio de uma visualização disposta em um plano hiperbólico que, para ser gerada, faz uso do desenho radial de árvores, onde a estrutura de uma árvore é mapeada em uma representação gráfica.

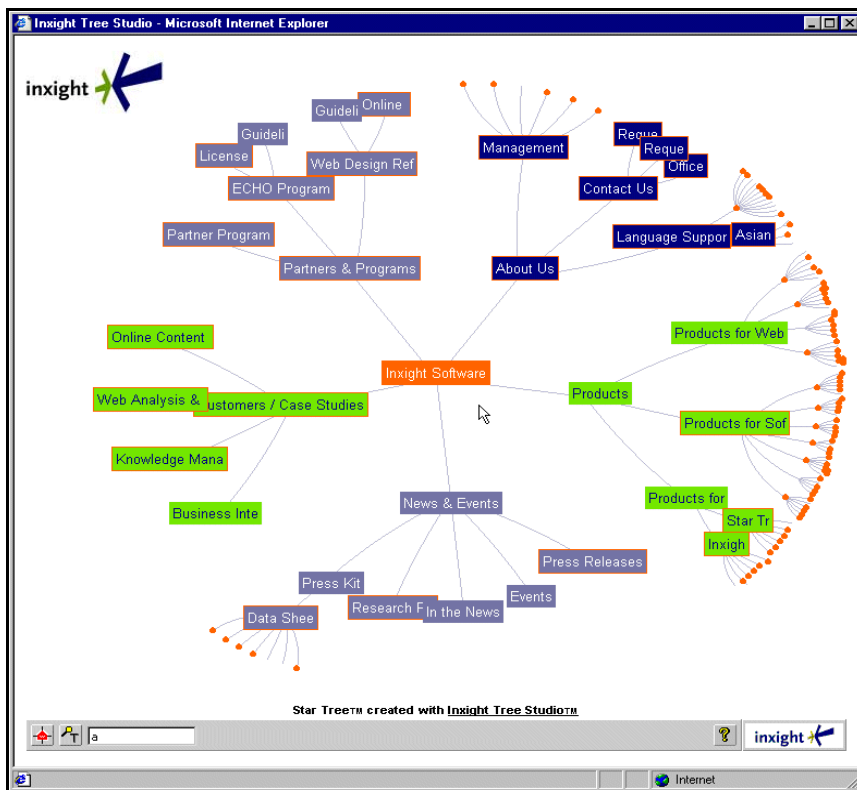
Na técnica browser hiperbólico, os conceitos estão organizados por hierarquias (categorias) tendo um assunto mais geral centralizado e, dentro deste, assuntos específicos relacionados a ele, por meio de vértices em uma espécie de árvore. Na estrutura, dentro de cada assunto específico estão dispostos os termos que se ligam a ele, por meio de ramificações que condensam os assuntos cada vez mais específicos, sendo possível, através desta estrutura, visualizar desde os assuntos mais abrangentes até aqueles menos inclusivos dentro da árvore, possibilitando a compactação das informações à medida que se aproxima da periferia do desenho.

Assim, essa técnica consegue disponibilizar cerca de dez vezes mais vértices de uma árvore do que utilizando uma visualização no plano cartesiano, dessa maneira, tem-se a possibilidade de representar uma grande quantidade de categorias em um espaço visual limitado. A navegação por meio do browser hiperbólico tende a tornar-se mais efetiva, podendo ser realizada através de simples movimentos do mouse, selecionando o vértice de maior interesse e arrastando-o para o centro para ter o foco desejado (NASCIMENTO; FERREIRA, 2005).

A estrutura do browser hiperbólico pode ser visualizada por meio da figura 9, que apresenta as informações sobre o *software* Inxight¹⁰, cuja função é a construção de mapas hiperbólicos para representar informações.

¹⁰ <http://inxight-software-inc.software.informer.com/>

Figura 9 - Browser Hiperbólico do software Inxight



Fonte: Xiang et al. (2005, p.74)

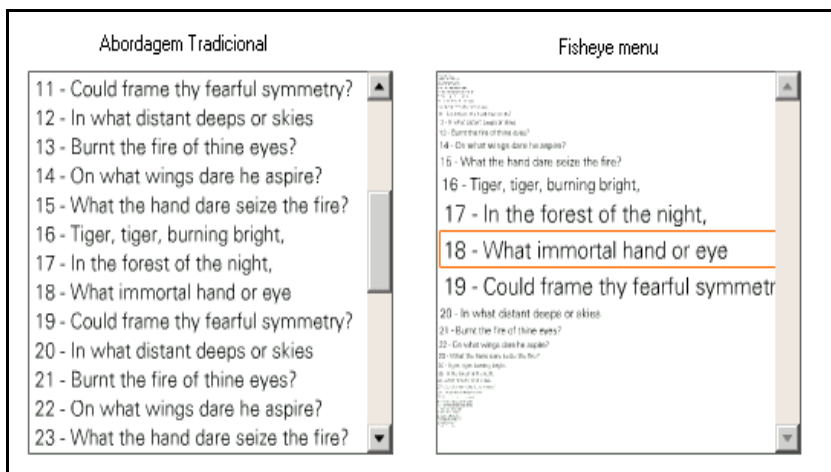
A partir da figura 9, percebe-se que o browser hiperbólico faz uso, juntamente com o desenho radial de árvores, de um importante princípio - o **foco+contexto**, cujo objetivo é apresentar uma visão geral dos dados a serem visualizados, destacando uma área específica de interesse (foco), por meio da ampliação suave da mesma enquanto há uma compactação do restante da imagem (NASCIMENTO; FERREIRA, 2005). Nas palavras de Xavier (2009), o funcionamento do foco+contexto consiste na distorção de

uma representação visual, tornando mais visível o foco e minimizando o contexto. Devido a sua funcionalidade e efetividade, o princípio foco+contexto é largamente empregado nas técnicas de VI, os principais exemplos de técnicas que utilizam o foco+contexto são: *Fish-eye*, *Perspective Wall* e *Table Lens*.

- *Fish-eye*

A técnica *fish-eye* é conhecida por produzir, como o próprio nome diz, um efeito semelhante à lente de aumento de um olho de peixe, possibilitando uma visão detalhada de uma região de interesse através de uma taxa maior de ampliação no centro dessa região e decrescente no sentido da periferia da imagem (NASCIMENTO; FERREIRA, 2011). Essa técnica, de acordo com Estivalet (2000), é útil quando se quer uma visão mais detalhada de uma determinada região sem perder o contexto da visualização dos arredores dessa região de interesse.

A figura 10 apresenta um exemplo de abordagem do *fish-eye*, chamada *fish-eye menu*. Na imagem à esquerda percebe-se uma abordagem tradicional de menu, onde uma barra de rolagem tem que ser utilizada para permitir acesso a todos os ítems do menu. Já à direita, tem-se uma visualização que emprega o conceito de *fish-eye*, onde todos os itens de uma região são apresentados simultaneamente, por meio de uma ampliação dessa região, sendo permitida a movimentação do mouse sobre a região do foco, causando a ampliação das opções próximas a ela.

Figura 10 – Fish-eye menu

Fonte: Nascimento e Ferreira (2005, p. 1295)

- Perspective Wall

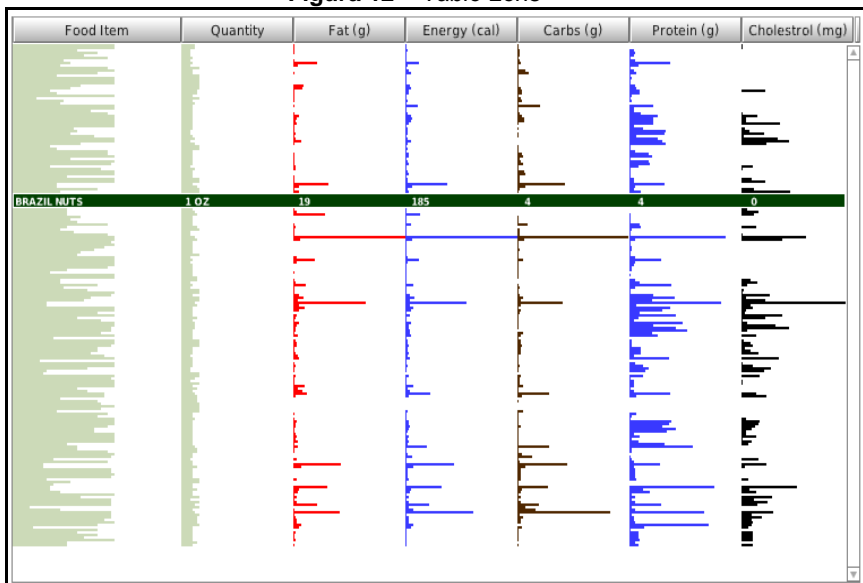
Essa técnica é bastante similar a *fish-eye*, pois provê um aumento uniforme em uma região regular de interesse enquanto o contexto global da estrutura é mantido. A *perspective wall* busca, especificamente, integrar o detalhe e o contexto de uma estrutura linear em uma única visualização, essa afirmação é compartilhada por Mackinlay, Robertson e Card (1991), ao mencionarem que a *perspective wall* integra uma visão detalhada e contextual para apoiar a visualização dos espaços de informação linearmente estruturados.

Nessa técnica as informações são representadas como um retângulo horizontal e a ideia é dispor as informações nesse retângulo, o qual é dobrado para trás nos limites à direita e à esquerda (definido pela área de interesse)

- *Table Lens*

Consiste em uma técnica utilizada para o entendimento de grandes volumes de dados numéricos e categóricos multidimensionais. Essa técnica permite a visualização de uma tabela, onde os dados (linhas) de interesse aparecem expandidos e os demais itens de forma compactada. As linhas da tabela são vistas como linhas de pixels formadas por gráficos de barras dos valores de seus atributos. A técnica possibilita interação do usuário para focalizar linhas ou colunas, o que causa uma ampliação delas enquanto mantém compactado o contexto ao seu redor (NASCIMENTO; FERREIRA, 2005). A figura 12, a seguir, traz um exemplo da técnica *table lens*.

Figura 12 – *Table Lens*



Fonte: Nascimento e Ferreira (2005, p.1297)

É relevante destacar que as técnicas *Table Lens* e *Perspective Wall*, citadas anteriormente, podem ser utilizadas com mais precisão para visualização dos resultados retornados por meio de uma busca. Essas técnicas, devido a sua alta dimensionalidade, podem representar uma baixa efetividade em determinados contextos, por isso precisam ser empregadas com cautela.

Além das técnicas mencionadas até o momento, se faz importante abordar os modelos de interface de VI, que funcionam, também, como ferramentas potencializadoras para representação, localização e recuperação da informação. Esses modelos são comumente empregados como interfaces visuais em sistemas de informação na web e serão discutidos em seguida.

4.3 Modelos de Interface de Visualização da Informação

Os modelos de interface de VI podem ser compreendidos como mecanismos interativos para representação e recuperação da informação e se configuram como sistemas de navegação que, de acordo com Xavier (2009), são cruciais para que os usuários possam explorar espaços informacionais visuais.

Assim, os modelos de interface se colocam promissores em tornar sistemas de busca mais acessíveis por meio de uma estrutura gráfica, que permite navegação interativa. A esse respeito, Eick (2001 apud XAVIER, 2009) acrescenta que esse é o foco da VI, a criação de interfaces

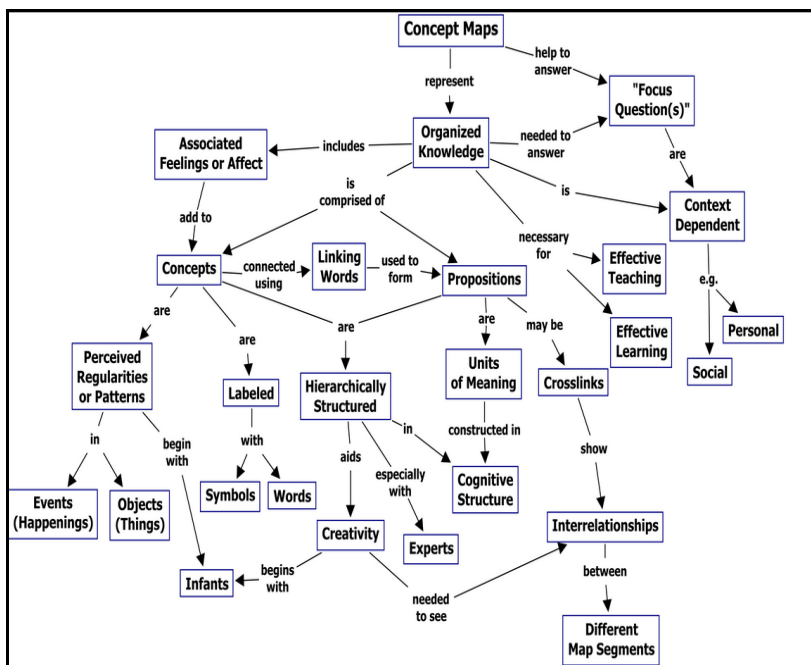
visualmente ricas para ajudar o usuário a compreender e navegar através de espaços de informações complexos.

Alguns importantes modelos de interface, presentes na literatura de Ciência da Informação, e que são amplamente utilizados como interfaces gráficas para SRIs, serão discutidos a seguir.

a) Mapa Conceitual

Os mapas conceituais são ferramentas usadas para representar graficamente informações e, para isso, fazem uso de diagramas que permitem uma melhor visualização acerca de como estão estruturadas as informações e como se relacionam entre si. O mapa conceitual possui diversas abordagens e pode ser empregado na representação, organização, visualização e recuperação de informações; sua utilização, como afirma Lima (2004), permite entender estruturas de informação de forma facilitada, por meio de uma interface atrativa, interativa e de fácil uso.

O mapa conceitual, de acordo com Novak e Cañas (2008), é formado por conceitos, geralmente localizados em círculos ou retângulos de algum tamanho, e por relacionamentos entre esses conceitos, que são indicados por uma linha que os conecta, formando proposições. Palavras sobre as linhas especificam as relações entre esses conceitos. Essa estrutura pode ser observada na figura a seguir, que apresenta um mapa conceitual explicando de forma gráfica o que são mapas conceituais.

Figura 13 - Mapa Conceitual sobre Mapas Conceituais

Fonte: Novak e Cañas (2008, p.1)

As duas importantes características do mapa conceitual são: 1- a representação dos conceitos de forma hierárquica, com o conceito mais geral no início do mapa e depois os mais específicos arranjados hierarquicamente; 2- a existência de referências cruzadas, que permitem verificar como está representada a relação entre os conceitos, ajudando a perceber como um conceito em um domínio do conhecimento está relacionado a um conceito em outro domínio representado no mapa (NOVAK; CAÑAS, 2008).

Assim, a representação gráfica e concisa de informações através de mapas conceituais e a possibilidade

de ligação entre conceitos, representados por termos ou palavras-chave, resultam em um ambiente ideal para se definir um sistema de navegação, no qual os usuários podem encontrar a informação que buscam, bem como navegar pelo sistema pesquisando outros temas (LIMA, 2004). Nessa linha de pensamento, Cañas et al. (2005) destacam, ainda, que quando os conceitos que representam informações e as ligações entre eles são devidamente estabelecidas, os mapas, enquanto interfaces gráficas, podem vir a ser considerados poderosas ferramentas para descobrimento de informações.

b) Cartografia Temática

Segundo Eliel (2007), a cartografia temática consiste em um mecanismo de representação das temáticas inscritas em recursos informacionais, tais como livros, periódicos, teses etc. Pode ser entendida como mapas gerados a partir de dados bibliográficos, proporcionando uma visualização gráfica desses dados, oferecendo uma visão global de conjuntos de informações e evidenciando relações e estruturas entre elas.

Existem dois tipos de cartografia temática: 1- a que retrata o espaço de documentos (mostra a relação de similaridade entre documentos e revela temas principais tratados por grupos de documentos); 2- a que retrata o espaço dos termos (mostra a relação de coocorrência entre termos). Assim, a cartografia temática pode ser empregada tanto para agrupar documentos de acordo com assuntos semelhantes como para estabelecer relações de coocorrência e proximidade entre documentos, facilitando, em ambos os casos, a visualização e compreensão de como estão

organizados os conteúdos, e possibilitando uma navegação direcionada para interesses específicos.

A utilidade da visualização de informações, por meio de mapas ou cartografia temática, baseia-se em estudos sobre a percepção que mostram que o ser humano tem primeiro uma percepção global de uma cena antes de atentar para os detalhes (KOBASHI; SANTOS, 2008). Tais estudos abriram caminhos para explorar as características da percepção global e, conseqüentemente, para aplicá-las aos sistemas de informação, particularmente, nos aspectos relacionados à recuperação de informação, tais como:

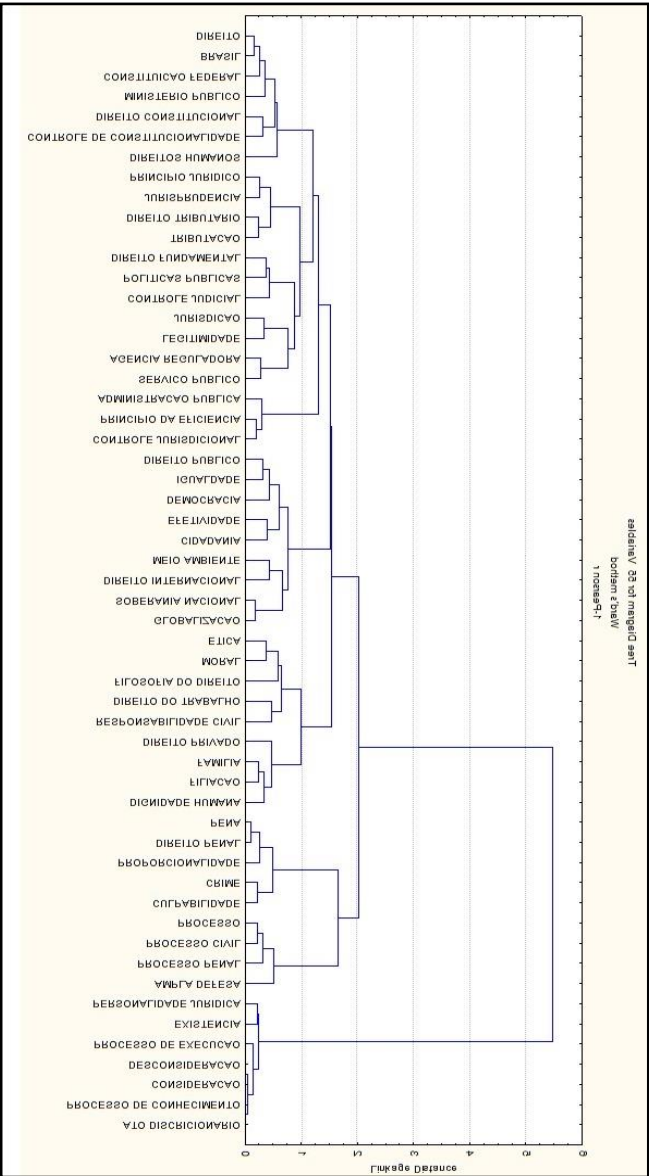
- Exploração rápida de conjuntos de informações desconhecidas;
- Evidenciação de relações e estruturas nas informações;
- Fornecimento de alternativas de acesso a informações pertinentes;
- Classificação interativa de informação.

Para melhor compreender a estrutura da cartografia temática, serão apresentadas a seguir duas figuras. Essas figuras revelam os termos que representam as temáticas das teses e dissertações do programa de pós-graduação em Direito da Universidade Federal de Pernambuco (UFPE).

A figura 14 foi criada a partir das relações de coocorrência entre os termos e traz os descritores por hierarquia em um dendrograma, permitindo visualizar as relações de proximidade temática. Já a visualização da figura 15 apresenta, através da escala multidimensional (MDS),

uma estrutura não hierárquica dos descritores, mas com relações associativas de proximidade entre as temáticas.

Figura 14 - Cartografia temática para o PPG em Direito, baseada em classificação hierárquica de descritores- DENDROGRAMA



Fonte: Vieira (2011, p. 40)

c) Nuvem de Tags

Nuvem de *tags* ou nuvem de etiquetas (do inglês *Tag Cloud*) consiste em uma interface de VI usada, geralmente, por websites e blogs. Essa interface é composta por palavras-chave dispostas em uma lista visual com o objetivo de representar itens de conteúdo de um sistema. A nuvem de *tags* é, também, onde se concentram as *tags* mais utilizadas e/ou vistas no momento, ou seja, as palavras mais cobçadas pelos inúmeros acessos dos usuários, e, nesse sentido, sua função, como aponta Guedes e Souza (2008), é permitir a rápida identificação dos temas mais destacados pelos usuários.

Como mencionado acima, o modelo nuvem de *tags* é comumente utilizado no ambiente web, ambiente em que, conforme argumentam Nielsen e Loranger (2007 apud MOREIRA; ROMÃO, 2009), são adotadas estratégias de busca de informações para economizar tempo, ou seja, nesse ambiente os usuários não tendem a buscar informações de uma maneira linear, em vez disso, contam com pistas visuais que fornecem uma indicação mais consistente. Nesse contexto, as palavras da nuvem de *tags* podem funcionar como pistas para os usuários, pois tem a função de facilitar a navegação, dando orientações e indicando por onde esses usuários poderiam caminhar para localização de informações de seu interesse.

O acesso ao conteúdo informacional nesse modelo de interface é realizado, também, de maneira simples, pois as *tags* disponibilizadas na nuvem funcionam como links que

levam a coleções de itens relacionados, por meio de um clique do mouse em uma tag é possível ter acesso a uma página onde há vários links para documentos que foram indexados por essa tag (VIEIRA, 2011). Guedes e Souza (2008) acrescentam, nesse sentido, que uma vez que o usuário clica sobre uma tag, ele obtém uma lista ordenada de recursos descritos por aquela palavra, como também uma lista de palavras relacionadas com a primeira, possibilitando continuar a navegação. Dessa forma, a nuvem de *tags*, por meio de sua estrutura amigável, facilita a navegação, localização e recuperação de informações pelo usuário-navegador.

Como exemplo desse modelo de interface, tem-se a figura 16 a seguir, a qual representa as *tags* mais populares do *Flickr*¹¹, que é um site da web de hospedagem e partilha de documentos gráficos como imagens fotográficas, desenhos e ilustrações.

¹¹ <http://www.flickr.com/photos/tags/>. O site de compartilhamento de imagens *Flickr* foi a primeira página da web com ampla divulgação a usar a nuvem de etiquetas (GUEDES; SOUZA, 2008).

Figura 16 - Nuvem de Tags

Fonte: <http://www.flickr.com/photos/tags/>

Na figura 16 algumas *tags* como *instagramapp* e *square squareformat* possuem um destaque maior em relação a outras dentro da nuvem, esse recurso é utilizado para representar os assuntos que estão em foco no sistema, ou seja, os de maior ocorrência no momento. Desse modo, a nuvem de *tags* evidencia os assuntos mais acessados do sistema proporcionalmente pelo tamanho da fonte de suas palavras.

Com relação ao conjunto de palavras que compõem a nuvem de *tags*, é relevante mencionar que a escolha dessas

palavras, em alguns casos, tem sua origem em um processo de indexação utilizando folksonomia. Folksonomia, a saber, é a prática de indexação colaborativa de informações em meio digital, realizada pelos usuários que utilizam o seu próprio vocabulário (linguagem natural) de forma livre, sem preocupação com padronização, para indexar páginas ou informações de seu interesse contidas na web (MOREIRA; ROMÃO, 2009). A folksonomia visa, portanto, o compartilhamento das informações entre os leitores nesse ambiente.

Nessa perspectiva, muitas vezes, os assuntos mais importantes de uma página acabam sendo indexados através das *tags* criadas pelos próprios internautas e expostas no formato de nuvem, podendo funcionar até como uma consulta pronta para os próximos navegadores (MOREIRA; ROMÃO, 2009).

Após análise dos aspectos que envolvem a VI, suas técnicas e modelos de interface, será abordado, na seção seguinte, o método utilizado para o desenvolvimento da presente pesquisa.

5 MÉTODO

Nesta seção será apresentado o método que caracteriza a pesquisa bem como o percurso metodológico utilizado para sua realização.

5.1 Tipo de Pesquisa

Esta pesquisa se caracteriza como exploratória, uma vez que, para o alcance dos objetivos propostos, buscou-se levantar e explorar referencial teórico acerca da organização e representação visual de informações como alternativa para potencializar o acesso e a recuperação do conteúdo informacional armazenado em Sistemas de Recuperação de Informação.

Assim, por meio da pesquisa exploratória, que, de acordo com Gil (2009, p. 27) “[...] tem como principal finalidade desenvolver, esclarecer [...] conceitos e ideias, tendo em vista a formulação de problemas mais precisos [...]”, buscou-se proporcionar uma maior familiaridade com as áreas abordadas.

Com o intuito de possibilitar a formação de um referencial teórico acerca das temáticas aqui estudadas, a pesquisa se caracteriza, também, como bibliográfica. Pois, a partir da pesquisa bibliográfica obtém-se o embasamento científico para melhor compreensão e análise de cada área.

5.2 Etapas da Pesquisa

Para o cumprimento dos objetivos propostos na pesquisa, alguns procedimentos metodológicos são necessários. Desse modo, tais procedimentos serão descritos a seguir a partir das etapas percorridas no desenvolvimento desta pesquisa.

5.2.1 Levantamento Bibliográfico

O levantamento bibliográfico baseou-se em literatura voltada para as temáticas aqui abordadas, com o intuito de fundamentar a pesquisa, e, para isso, foram analisadas referências teóricas extraídas de artigos, teses, dissertações e livros, presentes na literatura nacional e internacional, que versem sobre Organização da Informação, Visualização da Informação e Sistemas de Recuperação de Informação.

É importante mencionar que os artigos utilizados foram encontrados, em sua maioria, em periódicos e eventos da área de Ciência da Informação, com exceção de alguns trabalhos advindos da área da Tecnologia e da Ciência da Computação, que foram fundamentais para o desenvolvimento da pesquisa, principalmente para obtenção de informações mais específicas acerca das técnicas e modelos de VI.

As buscas por artigos foram realizadas, também, nas Bases de Dados: BRAPCI e SCIELO; e no portal de periódicos da CAPES, onde, por meio dos termos que versem

sobre as temáticas aqui abordadas: Organização da Informação, Visualização da Informação e Sistemas de Recuperação de Informação, foram realizadas as pesquisas por artigos pertinentes.

Para isso, alguns critérios foram adotados, destacando, especificamente:

- BRAPCI – entre os artigos localizados na busca, foram escolhidos aqueles em que as temáticas estivessem presente no título, no resumo e, principalmente, nas palavras-chave, o que facilitou na identificação de documentos que realmente abordassem os temas de interesse;
- SCIELO – de maneira semelhante, a escolha por artigos pertinentes foi realizada adotando como critério aqueles artigos em que os termos, anteriormente mencionados, constassem no título e no resumo dos documentos apresentados. Além disso, as buscas foram realizadas utilizando o refinamento de resultados: áreas temáticas, sendo escolhida a área Information Science and Library Science. A possibilidade de pesquisa por meio dessa especificação permitiu a recuperação de artigos dentro da área de interesse.
- Portal de periódicos da CAPES – Aqui, entre as opções de buscas existentes, foram realizadas buscas por assunto, contando, também, com o auxílio do refinamento de resultados por áreas temáticas. Dessa maneira, obteve-se uma relação de artigos, presentes na base Directory of Open Access Journals (DOAJ), dentre os quais foram escolhidos aqueles que apresentaram as temáticas de interesse no título e no resumo.

Todas as pesquisas foram realizadas sem restrição de data, entretanto, as referências utilizadas neste trabalho datam do período de 1971 a 2013.

5.2.2 Descrição de técnicas e modelos de interface de VI

Diante da existência de diversas técnicas e modelos de interface de visualização da informação, optou-se, nesta pesquisa, por analisar aqueles que melhor se adéquam aos SRIs, ou seja, que tenham aplicações práticas na construção de interfaces para sistemas que lidam com a recuperação da informação, possibilitando a representação e visualização da estrutura da coleção de documentos. Nessa perspectiva, quanto às técnicas, foram analisadas:

- Desenho de Grafos;
- Desenho de Árvores;
- Browser Hiperbólico.

A partir da técnica Browser Hiperbólico, que faz uso do princípio foco+contexto, abordou-se, também, as principais técnicas que adotam esse princípio, a saber:

- Fish-eye;
- Perspective Wall;
- Table Lens.

Essas últimas são caracterizadas por sua aplicação na visualização de resultados retornados por meio de uma busca.

Quanto aos modelos de VI, conforme mencionado anteriormente, foram analisados os seguintes:

- Mapa Conceitual;
- Cartografia Temática;
- Nuvem de *Tags*.

Tais modelos são amplamente utilizados como interfaces gráficas para SRIs.

5.2.3 Análise de sites de SRIs

Nesta etapa da pesquisa, foi feito um levantamento de SRIs na Web que adotam visualizações na interface com o usuário. Os sites foram abordados com o propósito de identificar as técnicas e modelos empregados na concepção de suas interfaces, de forma a permitir uma visualização prática da utilização de interfaces em SRIs.

Nesse contexto, apresenta-se uma análise de como deveriam estar organizadas e estruturadas as informações e quais instrumentos seriam necessários para tornar possível a aplicação eficaz de cada técnica e modelo de VI nos sistemas, uma vez que, para que as interfaces visuais possam cumprir seu papel, facilitando o acesso e recuperação de informação, os sistemas que as adotam necessitam ter suas informações devidamente organizadas, conforme exposto ao longo da pesquisa.

Para a análise, foram abordados os seguintes sites de SRIs e suas interfaces visuais:

- AquaBrowser da biblioteca da Universidade de Edimburgo - o catálogo on line da biblioteca faz uso de interface baseada na técnica desenho de grafos;
- Mecanismo de busca YIPPY - faz uso do desenho de árvores de diretórios para organizar os documentos retornados para uma consulta realizada em sua base de documentos;
- Infoteca da Embrapa – faz uso da técnica Browser Hiperbólico para permitir acesso a informações sobre tecnologias produzidas pela Embrapa;
- Website do IHMC Camp Tools - utiliza o mapa conceitual como modelo de interface para descrever o programa Cmap Tools;
- WEBSOM - método para construção de mapas visuais – utiliza a cartografia temática como um método para gerar mapas visuais de assuntos;
- Sistema de busca da BRAPCI – faz uso da nuvem de *tags* como interface visual para auxiliar o usuário, sugerindo termos de busca.

Cada um dos sistemas foi escolhido por fazer uso de uma técnica ou modelo de interface visual anteriormente abordado, e por ter como finalidade a representação e recuperação de documentos como registros de informação.

Em cada site foi analisado, especificamente:

- Qual técnica ou modelo de interface de VI o sistema adota em sua estrutura?
- Qual é forma de busca oferecida pelo sistema?

- Como o sistema apresenta os resultados de uma busca?
- As visualizações adotadas no sistema são expressivas e efetivas?
- Que vantagens as interfaces visuais trazem para os usuários que interagem com elas?
- Como devem estar organizadas as informações na base de dados do sistema para que a interface visual possa ser adotada?
- Quais instrumentos são necessários para tornar possível a construção da interface visual?

6 SISTEMAS QUE FAZEM USO DA VISUALIZAÇÃO DA INFORMAÇÃO

A VI, por meio de suas técnicas e modelos de interface, pode ser considerada uma poderosa aliada para os sistemas que se propõem a oferecer uma interface amigável e interativa para acesso, localização e recuperação das informações por eles armazenadas. Pois, como afirmam Torres Ponjuan e Ponjuan Dante (2010), o uso de recursos visuais para representação de informação possibilita o rápido acesso, localização e compreensão do que está sendo transmitido.

Nessa perspectiva, Zhang (2008) também argumenta que as técnicas de VI, que possuem a capacidade de demonstrar relações entre conceitos em um ambiente visual, transparente e interativo, configuram-se em importante alternativa para tornar os SRIs mais efetivos em suas atividades, pois oferecem uma forma única para revelar informações ocultas, bem como para permitir aos usuários a busca por informações, por meio de uma representação visual.

Nesse contexto, busca-se, a seguir, abordar alguns sites de SRIs que utilizam interfaces visuais em sua estrutura, demonstrando, de forma prática, como a aplicação das técnicas e modelos de interface de VI pode auxiliar na localização e recuperação de informações. A partir da análise desses sistemas, é possível observar, também, as vantagens que as interfaces visuais trazem para os usuários que acessam e interagem com elas.

Outro aspecto a ser analisado nas interfaces visuais dos SRIs diz respeito à expressividade e efetividade (atributos abordados na subseção 4.1). Com base nesses atributos, pode-se identificar se a visualização dos sistemas apresenta bom desempenho na sua interação com o usuário.

Por outro lado, os sistemas são analisados sob a perspectiva da organização da informação, ou seja, buscou-se compreender como deveriam estar organizadas as informações nas bases de dados dos SRIs e quais instrumentos seriam necessários para tornar possível a construção de uma interface visual para esses sistemas, tendo em vista que a organização prévia das informações, bem como a adoção de instrumentos adequados para isso, são elementos necessários para a representação de informações por meio de estruturas de VI.

Dessa forma, foram abordados os seguintes sites de SRIs: AquaBrowser, da biblioteca da Universidade de Edimburgo; Infoteca-e, da Embrapa; Website do IHMC *CmapTools*; WEBSOM - método para construção de mapas visuais; Mecanismo de busca YIPPI; Sistema de busca da BRAPCI; escolhidos, em especial, por fazerem uso das técnicas e modelos de VI anteriormente abordados.

6.1 AquaBrowser (Biblioteca da Universidade de Edimburgo)

A Universidade de Edimburgo, considerada uma das mais antigas da Escócia e uma das maiores e mais

prestigiosas universidades do Reino Unido, disponibiliza em seu site¹², além de informações sobre ensino, pesquisa e inovação, o sistema de biblioteca da universidade. Esse sistema de biblioteca oferece, entre outros serviços, a possibilidade de acesso ao catálogo da biblioteca, que contém mais de um milhão de itens compostos por coleções, livros e revistas em formato impresso e eletrônico.

Para acesso ao conteúdo desse catálogo, o sistema oferece duas alternativas: o catálogo clássico, que recupera registros por meio de palavras-chave, apresentando uma lista com os resultados; e o catálogo que faz uso do *AquaBrowser Library*¹³, oferecendo uma interface visual como alternativa simples e interativa para a busca por registros. O catálogo que faz uso da interface visual será, portanto, o foco da análise aqui realizada no site da biblioteca da Universidade de Edimburgo. Para ter acesso a esse catálogo, basta apenas localizar o link “*Aquabrowser catalogue search*”, na página da biblioteca¹⁴. Para ilustrar a interface do catálogo, apresenta-se a figura 17.

¹² <http://www.ed.ac.uk/home>

¹³ *AquaBrowser Library* fornece uma interface nova e moderna que encoraja o usuário a explorar o conteúdo da biblioteca por comportamentos de busca recompensadores. Sua metodologia oferece recursos que ajudam no descobrimento de resultados relevantes de maneira rápida e fácil, convidando o usuário a interagir e se envolver com o conteúdo da biblioteca.

¹⁴ <http://www.ed.ac.uk/schools-departments/information-services/library-museum-gallery>

Figura 17 – Desenho de grafos do AquaBrowser da Biblioteca da Universidade de Edimburgo



Fonte: <http://aquabrowser.lib.ed.ac.uk/>

O sistema da biblioteca, apresentado na figura 17, faz uso de interface visual baseada na técnica desenho de grafos para realizar buscas no catálogo. Assim, a partir de uma busca expressa na caixa de texto, o sistema gera uma

visualização apresentando os resultados para essa busca por meio de termos, representados por vértices, e das relações entre eles por arestas. Dessa forma, o termo utilizado na busca permanece no centro da figura enquanto os termos relacionados a ele estão ligados por meio das arestas. O tipo de relação estabelecida entre os termos é determinado pela cor dos vértices. Assim, como demonstrado na figura 17, tem-se:

- **Associações** - termos que, de alguma forma, estão relacionados ao termo pesquisado (como **model**, **geographic** e **application**, que possuem associação com o termo **cartography**);
- **Traduções** - diferentes idiomas em que o termo de busca é encontrado no acervo da biblioteca (como **cartografia** e **mapmaking**);
- **Variações de ortografia** - diferentes formas para o termo utilizado, permitindo, também, visualizar outras maneiras de soletrar uma palavra (como **cartographer**, **cartographic**). As variações de ortografia ajudam, ainda, na correção de erros de ortografia;
- **Trilha descoberta** - termo da pesquisa inicial e subsequentes termos selecionados ao longo da busca (como **visualization**, **tool**, **technique**).

Dessa maneira, por meio do termo utilizado na busca, o usuário tem a possibilidade de verificar outros termos que estão interligados ao inicial, bem como as possíveis representações do termo em outros idiomas e, ainda, verificar

as variações ortográficas, sugerindo termos semelhantes ou, até mesmo, correções, caso o termo pesquisado apresente erros de digitação. Além disso, os caminhos percorridos na visualização durante o processo de busca, ou seja, os termos selecionados durante a navegação, ficam registrados em azul, conforme visto anteriormente, para que o usuário tenha uma visão do caminho percorrido caso queira retornar ou situar a questão central.

Por meio da navegação, o usuário pode, também, ter acesso ao catálogo da biblioteca, pois, ao clicar nos termos representados no desenho de grafos, o sistema retorna uma lista de resultados referentes ao termo escolhido. Essa lista exibe resultados que podem ser classificados por relevância, ano, autor e título. O sistema permite, ainda, refinar os resultados da busca, pois oferece uma variedade de categorias, como formatos específicos, autor, assunto, linguagem, séries, períodos de tempo etc. Assim, o usuário tem a possibilidade de redirecionar o foco para informações ainda mais específicas, possibilitando, dessa forma, explorar profundamente o conteúdo do catálogo.

6.1.1 Vantagens oferecidas pela interface visual da biblioteca da universidade de Edimburgo

Diante do exposto acerca da interface de visualização baseada em desenho de grafos, pode-se perceber que essa visualização permite ao usuário navegar, de forma amigável; conhecer opções para refinar sua busca, incorporando termos a ela; facilitar na localização de informações precisas e,

também, contribuir para dedução de novos conhecimentos, já que, a partir dos relacionamentos sugeridos, o usuário tem a oportunidade de visualizar conteúdos diferentes que podem ser de seu interesse, mas que antes eram desconhecidos.

Outra vantagem oferecida por essa visualização baseia-se no fato de que é recorrente o usuário não ter bem definida sua questão de busca ao acessar um sistema. Sendo assim, o desenho de grafos pode ajudá-lo a encontrar o que procura, levando em conta que, mesmo que o termo pesquisado não seja exatamente o melhor termo para pesquisa, as relações e variações sugeridas pela visualização podem, em certos casos, auxiliar o usuário, direcionando-o para informações precisas.

Em síntese, é possível afirmar que a visualização gerada pelo desenho de grafos contribui para localização e recuperação de conteúdos pertinentes, por meio de uma interface visual e interativa.

6.1.2 Análise da interface visual quanto à expressividade e efetividade

A visualização adotada pela biblioteca da universidade de Edimburgo pode ser considerada expressiva, pois, além de exibir resultados centrados na pesquisa do usuário, apresenta as relações encontradas das palavras digitadas com outras no vocabulário do sistema. Assim, de maneira geral, a visualização busca atender ao usuário expressando

informações relevantes que estejam dentro do seu contexto de interesse, razão pela qual é considerada expressiva.

Além disso, a visualização da biblioteca da universidade de Edimburgo pode ser considerada efetiva, uma vez que facilita a compreensão das informações transmitidas e permite ao usuário explorar, de maneira amigável, a relação entre termos para selecionar aqueles que lhe sejam mais convenientes.

6.1.3 Análise dos elementos de OI necessários para o emprego de interface visual baseada no desenho de grafos

Para que a visualização baseada na técnica desenho de grafos possa exercer adequadamente sua função é importante que uma atenção maior seja dada quanto à organização prévia das informações a serem disponibilizadas, especificamente, quando se refere aos termos usados para representar os documentos, e as relações semânticas estabelecidas entre eles no sistema.

Tendo em vista que o foco da técnica desenho de grafos consiste na representação de conceitos (termos) e na visualização de relacionamentos estabelecidos entre eles (demonstrando as relações entre termos pesquisados pelo usuário com outros termos presentes no vocabulário do sistema), se faz necessário, então, a adoção de instrumentos de OI que estabeleçam e evidenciem as relações existentes entre conceitos.

Neste sentido, têm-se como instrumentos que realizam essa função os tesauros e a taxonomia. Os tesauros são capazes de organizar termos relacionados semântica e genericamente, e são considerados, conforme Carlan e Medeiros (2011), instrumentos de controle terminológico eficazes para a organização de conceitos e importantes ferramentas no tratamento e recuperação da informação. No mesmo contexto está inserida a taxonomia, como instrumento que permite agrupamento categorizado de assuntos e a construção de relacionamentos semânticos entre termos, sendo considerada, também, como menciona Vital (2012, p. 26), “uma metodologia facilitadora tanto da organização quanto da recuperação, representação e disponibilização do conteúdo em ambiente Web”.

Sendo assim, o auxílio dos instrumentos de OI para uma representação e organização adequada de termos pode permitir que a técnica desenho de grafos seja aplicada e que, por meio dela, o usuário tenha uma compreensão visual clara acerca das diversas associações e vertentes disponíveis para um termo de busca.

6.2 Mecanismo de busca YIPPY

O YIPPY é um mecanismo de metabusca, ou seja, um sistema de busca que envia solicitações de usuários para motores de busca e/ou bancos de dados com o intuito de recuperar conteúdos que atendam a essas solicitações. O diferencial do sistema YIPPY está no fato de que os resultados obtidos de uma busca são agregados em

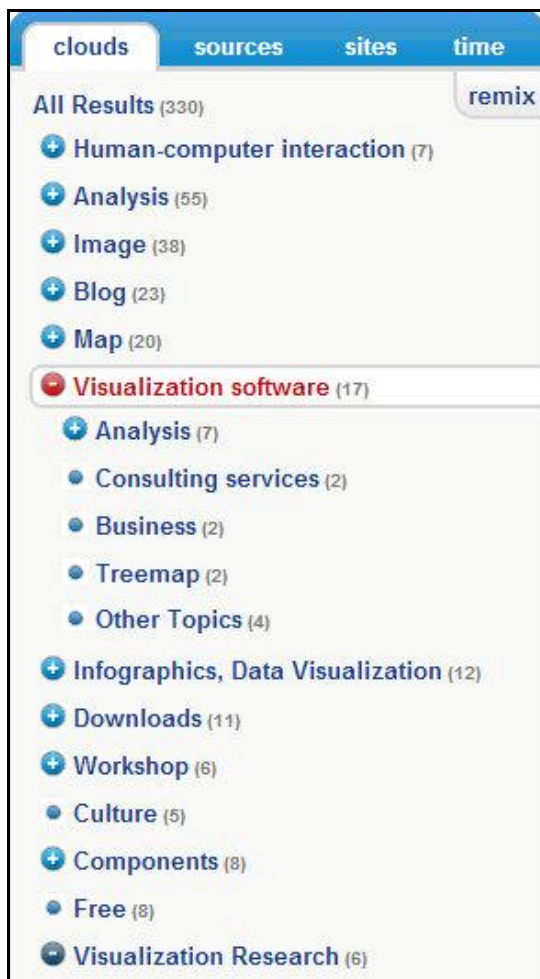
categorias e dispostos em uma lista ordenada, por meio de uma interface simples e intuitiva.

O mecanismo oferecido pelo YIPPY para organização dos itens recuperados será o ponto principal a ser analisado em seu site¹⁵, que disponibiliza uma interface visual baseada na técnica desenho de árvores de diretórios. Na interface disponibilizada pelo YIPPY os documentos retornados a partir de uma busca são agrupados de acordo com a semelhança entre eles, definida a partir da ocorrência de determinados termos (tópicos), representados por conteúdos comuns.

A figura 18, a seguir, ilustra a visualização gerada pelo YIPPY para apresentação dos resultados de uma busca realizada pelo termo *information visualization*.

¹⁵ <http://yippy.com/>

Figura 18 - Desenho de árvores do YIPPY



Fonte: <http://yippy.com/>

Para que a busca possa ser realizada, o YIPPY fornece uma caixa de texto em sua página inicial e, também, no topo de cada página de resultados. Dessa forma, a partir

da expressão de um termo, o sistema exibe os resultados por categorias representadas por tópicos e, também, por subtópicos, que funcionam como elementos de discriminação dos documentos dentro de determinado tópico.

Essas categorias podem ser visualizadas na figura 18, anteriormente apresentada, a qual ilustra uma busca realizada pelo termo *information visualization*, sendo possível observar, dentro dos resultados retornados para esse termo, o tópico *visualization software* e, relacionados a ele, os subtópicos *analysis* e *treemap*. Vale salientar que tópicos e subtópicos funcionam, também, como elementos de refinamento de busca.

A partir da visualização gerada com os resultados da busca, é possível ter acesso aos documentos, também de forma simples. Para isso, é necessário apenas selecionar um tópico e, em seguida, obtém-se, à direita, uma relação de itens com o conteúdo referente àquele tópico.


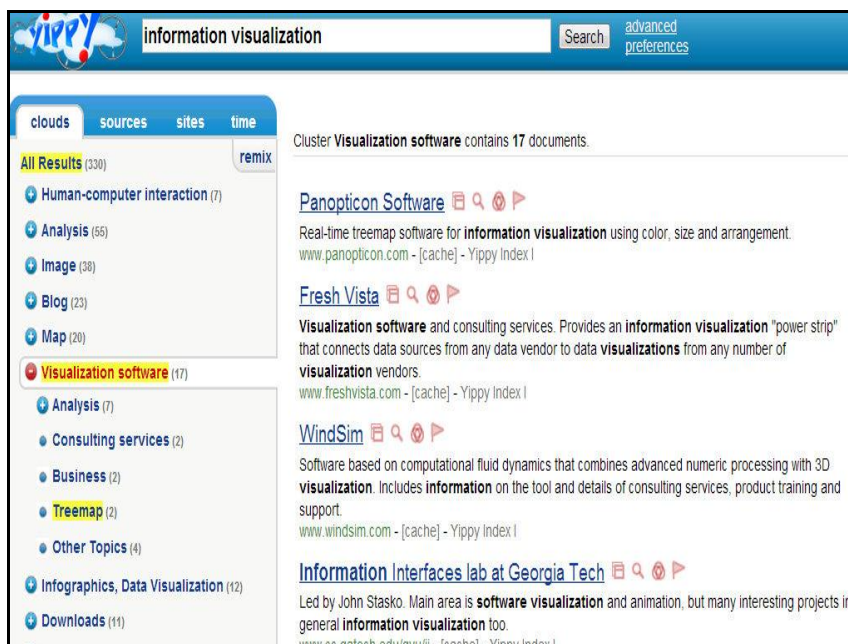
Entre os resultados retornados, é possível ainda destacar, na visualização onde estiver localizado, um item em questão, ou seja, na categoria em que um determinado documento está inserido; esse destaque é feito por meio de uma marcação na cor amarela, como pode ser visto na figura 19. Para isso, basta clicar no símbolo  ao lado de um item recuperado, para que a visualização destaque em que tópico está inserido o item e quais tópicos possuem documentos semelhantes a ele, permitindo, desse modo, a identificação de diferentes tópicos relacionados ao conteúdo selecionado.

Figura 19 – Resultado de uma busca no desenho de árvores do YIPPY



Fonte: <http://yippy.com/>

Como pode ser percebido na figura 19, ao lado de cada tópico/subtópico aparece, entre parênteses, uma numeração. Essa numeração corresponde à quantidade exata de documentos disponíveis para aquela categoria. Desse modo, o usuário tem a noção do número de documentos que tratam da temática especificada no tópico/subtópico antes mesmo de navegar pelos documentos.

6.2.1 Vantagens oferecidas pela interface visual do mecanismo de busca YIPPY

A visualização gerada pelo YIPPY fornece uma maneira amigável de busca. Na verdade, o que torna o YIPPY amigável é o que acontece na apresentação dos resultados da busca. Ao invés de disponibilizar uma longa lista com os resultados recuperados, o YIPPY oferece, a partir de uma interface visual, os assuntos agrupados em tópicos, minimizando a sobrecarga de informação sobre o usuário. Assim o usuário não necessita procurar o que deseja entre as inúmeras páginas de resultados, podendo visualizar os assuntos recuperados agrupados em única tela.

Além disso, a visualização do YIPPY permite explorar os resultados obtidos numa busca por meio de uma lista de tópicos ordenada alfabeticamente, o que aumenta a possibilidade de o usuário selecionar exatamente o que procura, ou descobrir relações entre tópicos e grupos de documentos.

Ao selecionar um item recuperado, o usuário tem, ainda, a possibilidade de visualizar em que tópico aquele item está inserido e se algum outro tópico pode conter itens associados ao assunto abordado. Assim, a partir do destaque em amarelo, o usuário pode refinar sua busca ou encontrar documentos inesperados dentro da própria visualização, o que facilita ainda mais a busca do usuário e torna a YIPPY uma poderosa ferramenta de pesquisa.

6.2.2 Análise da interface visual quanto à expressividade e efetividade

Acerca da visualização utilizada pelo site do YIPPY, pode-se argumentar que se trata de uma visualização expressiva, já que, na tentativa de apresentar para o usuário os resultados pertinentes a sua busca, a interface visual mostra apenas os tópicos que contêm os documentos referentes à temática de interesse do usuário, além de possibilitar explorar subtópicos semelhantes que permitem o direcionamento para assuntos específicos e mais precisos.

Quanto à efetividade, a visualização do YIPPY pode ser considerada efetiva, tendo em vista que sua estrutura não apresenta dificuldades para o usuário na localização de conteúdos de seu interesse, ao contrário, permite explorar a relação entre tópicos e grupos de documentos e selecionar aqueles mais relevantes, de forma simplificada. Além disso, informa o usuário quanto à quantidade de documentos presentes em cada tópico/subtópico.

6.2.3 Análise dos elementos de OI necessários para o emprego de interface visual baseada no desenho de árvores

Quando se refere à organização da informação necessária para que a técnica de VI desenho de árvores seja aplicada em sistemas que lidam com a recuperação de informação, é possível compreender que, para que isso ocorra, torna-se importante a adoção de um instrumento de

OI que auxilie na organização de conceitos, de maneira a possibilitar que a visualização cumpra o seu objetivo, ou seja, disponibilizar, de maneira organizada e simplificada, os resultados de uma busca.

Nesse contexto, pode-se citar a taxonomia como instrumento capaz de auxiliar a técnica desenho de árvores, visto que essa técnica, como foi observado na interface visual do YIPPY, busca apresentar os resultados de uma pesquisa por meio da categorização ou agrupamento de conteúdos semelhantes, evidenciando as relações hierárquicas existentes entre eles. É nesse contexto que se insere a taxonomia, como instrumento capaz de exercer essas funções, pois, segundo Vital (2012), as taxonomias buscam estabelecer categorias para a informação e apresentam relacionamentos semânticos entre termos.

A taxonomia, em si, é capaz de refletir a estrutura do desenho de árvores. Vital e Café (2011, p. 52) trazem uma rica contribuição a esse respeito ao destacar que:

Em relação à organização da informação recuperada do acervo, a taxonomia proporciona a categorização dos resultados de uma busca, facilitando a identificação das relações entre os assuntos. O mecanismo de busca do portal deve apresentar a saída dos resultados organizados nas categorias previamente definidas. Dessa forma, o usuário perceberá outros assuntos relacionados ao tema procurado que, muitas vezes, não lhe haviam ocorrido.

Dessa forma, a partir da adoção de uma taxonomia que permita a categorização de conceitos, a técnica desenho de árvores poderá exibir os resultados retornados de uma

busca agrupados em tópicos e subtópicos de acordo com as relações estabelecidas eles, auxiliando, assim, o usuário na localização de informações específicas, bem como no descobrimento de conteúdos relacionados ao que se procura, por meio de uma interface visual.

6.3 Infoteca-e da Embrapa

A Empresa Brasileira de Pesquisa Agropecuária (Embrapa) é uma instituição pública de Pesquisa, Desenvolvimento e Inovação (PD&I), vinculada ao Ministério da Agricultura, Pecuária e Abastecimento e tem como objetivo o desenvolvimento de tecnologias, conhecimentos e informações técnico-científicas voltadas para a agricultura e a pecuária.

A Embrapa tem como missão viabilizar soluções de pesquisa para a sustentabilidade da agricultura, em benefício da sociedade brasileira. Assim, a sua atuação junto à sociedade baseia-se numa estrutura organizacional composta de Unidades Administrativas, Unidades de Pesquisa e Unidades de Serviços. As unidades de pesquisa da Embrapa (centros de pesquisa) estão distribuídas em quase todos os estados do Brasil e suas ações de pesquisa têm abrangência nacional.

Para permitir o acesso às informações sobre tecnologias, as quais se relacionam às áreas de atuação de seus demais centros de pesquisa, a Embrapa disponibiliza em

seu site¹⁶ um link para o serviço Informação Tecnológica em Agricultura (Infoteca-e)¹⁷, que reúne conteúdos produzidos e editados na própria instituição com linguagem adaptada, de modo que seus usuários (produtores rurais, extensionistas, técnicos agrícolas, estudantes e professores de escolas rurais, cooperativas e outros segmentos da produção agrícola) possam assimilá-los com facilidade, apropriando-se, assim, de tecnologias geradas pela Embrapa.

A Infoteca-e da Embrapa apresenta como alternativa para fornecer acesso às informações sobre tecnologias produzidas pela empresa uma árvore hiperbólica, isto é, uma interface visual em forma de árvore espacial, baseada na técnica browser hiperbólico. Por meio da árvore hiperbólica é possível visualizar como está estruturado o acervo da Embrapa em termos de local, temática e ano. Essa interface visual permite o mapeamento e a visualização de informações, bem como das relações estabelecidas entre elas, configurando-se como uma facilitadora na compreensão, localização e descoberta de informações.

A árvore hiperbólica será, portanto, o ponto principal a ser analisado no site da Embrapa. Sendo assim, apresenta-se a figura 20, por meio da qual é possível visualizar, de forma gráfica, os registros das informações tecnológicas disponibilizadas pela empresa.

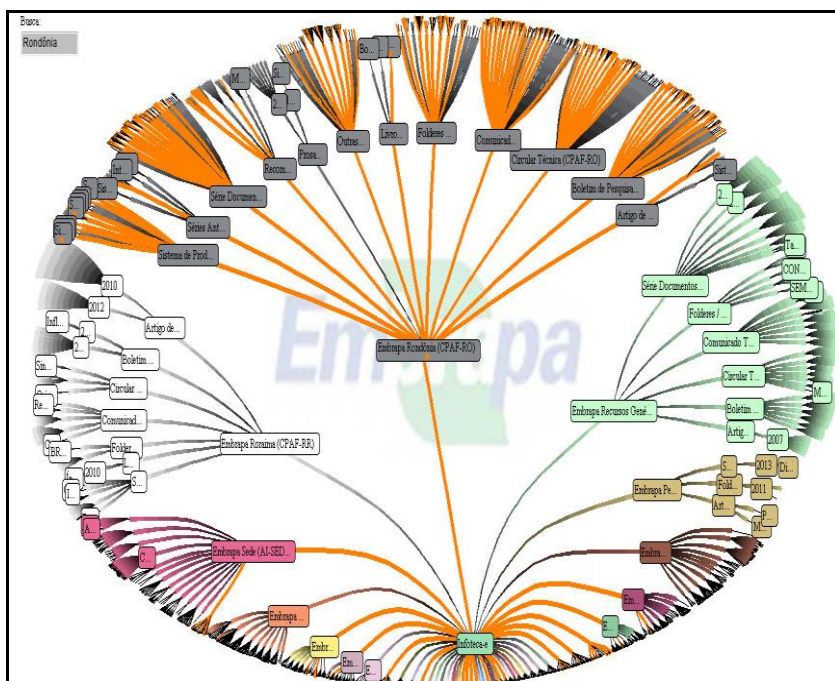
¹⁶ <http://www.embrapa.br/>

¹⁷ <http://www.infoteca.cnptia.embrapa.br/hypertree/hypertree.html>

meio das ramificações da árvore. Na figura 20, apresentada anteriormente, o vértice selecionado, que está em evidência no centro da figura, é o “Embrapa Pantanal (CPAP)”; na hierarquia, pode-se perceber as diversas ramificações pertencentes a esse vértice central, representando os assuntos cada vez mais específicos.

Outra maneira de realizar uma busca na árvore hiperbólica se dá por meio da caixa de texto, disposta ao lado esquerdo da interface visual. Por meio dela, pode-se expressar um termo de interesse, como pode ser visto na figura 21 e, em decorrência dessa busca, a interface evidencia (por meio de um destaque na cor laranja) os vértices e ramificações que contêm o termo pesquisado, direcionando o usuário para informações específicas, norteadas o caminho a ser percorrido.

Figura 21 - Busca realizada na árvore hiperbólica da Infoteca-e da Embrapa



Fonte: <http://www.infoteca.cnptia.embrapa.br/hypertree/hypertree.html>

Após a localização de um termo de interesse, pode-se ter acesso ao seu conteúdo, sendo necessário apenas um duplo clique sobre o vértice. A partir disso, o site remete o usuário aos links contendo os documentos completos referentes ao vértice selecionado.

6.3.1 Vantagens oferecidas pela interface visual da Infoteca-e da Embrapa

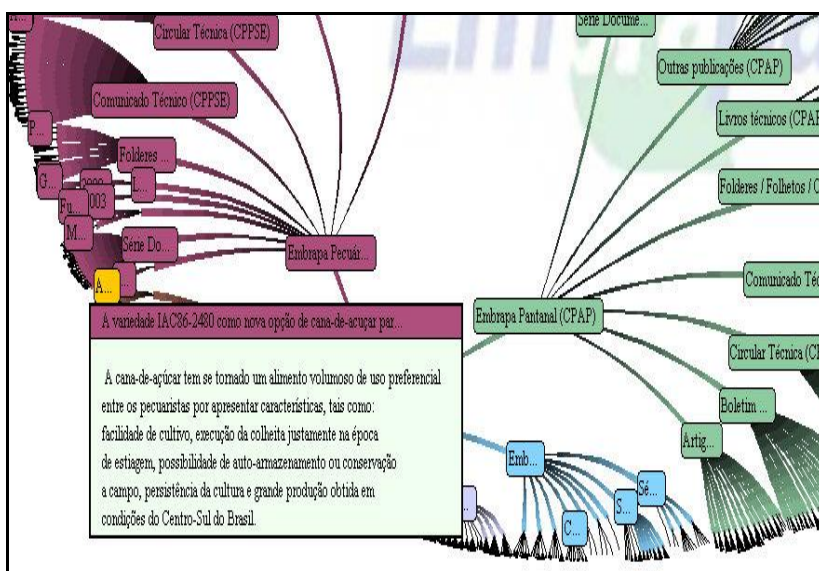
A visualização adotada pela Infoteca-e da Embrapa, baseada na técnica browser hiperbólico, pode ser considerada uma interface rica, por oferecer uma série de ganhos e otimizações aos usuários que dela fazem uso.

Uma dessas otimizações consiste no fato de que esse tipo de interface auxilia na disponibilização de conteúdos de forma organizada, facilitando na descoberta de informações em meio a uma grande quantidade de conteúdos disponíveis. Além disso, oferece uma visão geral dos conteúdos, auxiliando o usuário a compreender como estão estruturadas as informações; bem como oferece uma visão cada vez mais detalhada e mais aprofundada à medida que se navega por meio das hierarquias, tornando, dessa forma, mais simples a descoberta e a localização de conteúdos precisos.

Sob essa ótica, vale ressaltar que a interação com a interface em forma de árvore hiperbólica permite a identificação de conteúdos de maneira perceptiva e dinâmica; a partir de simples movimentos do mouse, pode-se selecionar os vértices de maior interesse, ajustando o foco desejado e percorrendo a interface na busca por conteúdos relevantes. Além das facilidades oferecidas pela navegação, conta-se, ainda, com a possibilidade da realização de buscas por meio da caixa de texto, o que acarreta em resultados destacados na visualização, por meio de cores e de trilhas visuais, tornando ainda mais fácil a localização de conteúdos específicos.

Vale mencionar, ainda, mais um ponto interessante que pode ser observado na interface da Embrapa: ao passar o mouse sobre os vértices da árvore, surgem informações descrevendo o conteúdo representado, conforme figura 22, permitindo ao usuário ter uma noção do que é tratado em cada vértice.

Figura 22 - Descrição do conteúdo do vértice



Fonte: <http://www.infoteca.cnptia.embrapa.br/hypertree/hypertree.html>

Assim, o usuário consegue ter uma ideia acerca do conteúdo abordado pelos vértices da árvore antes mesmo de abrir os documentos, despertando a atenção para informações de seu interesse, ou esclarecendo conteúdos que não sejam relevantes à sua procura.

6.3.2 *Análise da interface visual quanto à expressividade e efetividade*

Analizando a Infoteca-e da Embrapa quanto à expressividade e efetividade de sua visualização, pode-se argumentar que, apesar de disponibilizar em sua estrutura uma grande quantidade de informações, em que muitas acabam não sendo específicas ao interesse do usuário, esta visualização, ainda assim, pode ser considerada expressiva, pois oferece como alternativa uma caixa de texto, por meio da qual o usuário pode digitar termos e localizar informações de seu interesse, tendo as arestas que levam até os documentos que *casam* com a busca destacada na interface, facilitando a localização de informações específicas.

De maneira semelhante, a visualização da Infoteca-e pode ser considerada efetiva, já que permite a compreensão do que está sendo apresentado, não sendo de difícil entendimento na localização de informações.

No entanto, essa visualização, em específico, pode ter sua efetividade comprometida pelo fato de não permitir, durante a navegação, que o usuário encontre, de maneira simples, o vértice central (Infoteca) para retornar ao caminho inicialmente percorrido, tendo em vista que, à medida que se aprofunda nas ramificações, o foco da visualização vai sendo redirecionado e o ponto inicial fica encoberto na periferia da visualização, sendo necessária uma atenção maior quanto à aresta a ser selecionada para retornar ao ponto inicial.

6.3.3 Análise dos elementos de OI necessários para o emprego de interface visual baseada no browser hiperbólico

Para tornar possível que a técnica browser hiperbólico seja aplicada por um SRI e exerça com eficácia o seu propósito, trazendo benefícios na busca por informações, se faz necessário analisar que estrutura organizacional deveria ser adotada pelo sistema, ou seja, de que maneira o conteúdo desse sistema deveria estar organizado para dar suporte à construção da visualização.

Sob essa perspectiva, é importante ressaltar que uma árvore hiperbólica (interface gerada por meio da técnica browser hiperbólico) expressa seus conceitos (vértices) por meio de hierarquias (categorias). Sendo assim, para que essas hierarquias possam ser representadas em uma interface visual, elas precisam ser organizadas adequadamente, o que remete à necessidade de um instrumento de OI que permita uma organização baseada no estabelecimento de relações hierárquicas entre conceitos.

Nesse contexto, se insere a taxonomia como instrumento voltado à organização de informações em ambientes específicos e que, como afirma Vital (2012), é caracterizado pela sua estrutura hierárquica, capaz de organizar termos em categorias da mais genérica a mais específica, estabelecendo relações entre conceitos.

Dessa forma, tem-se na adoção de uma taxonomia uma alternativa para organização e representação categorizada de conceitos, que poderá auxiliar SRIs a construir uma interface baseada na técnica browser

hiperbólico, com o intuito de disponibilizar conteúdos de forma gráfica, evidenciando as relações entre eles e facilitando a recuperação eficaz de conteúdos.

6.4 Web site do IHMC CmapTools

O Institute for Human & Machine Cognition (IHMC) é um instituto de pesquisa sem fins lucrativos do Sistema Universitário da Flórida. O IHMC é formado por profissionais que investigam temas relacionados ao sistema cognitivo e à computação centrada no ser humano. Suas pesquisas colaboram para o desenvolvimento da ciência e da tecnologia, sendo de grande valor para a sociedade.

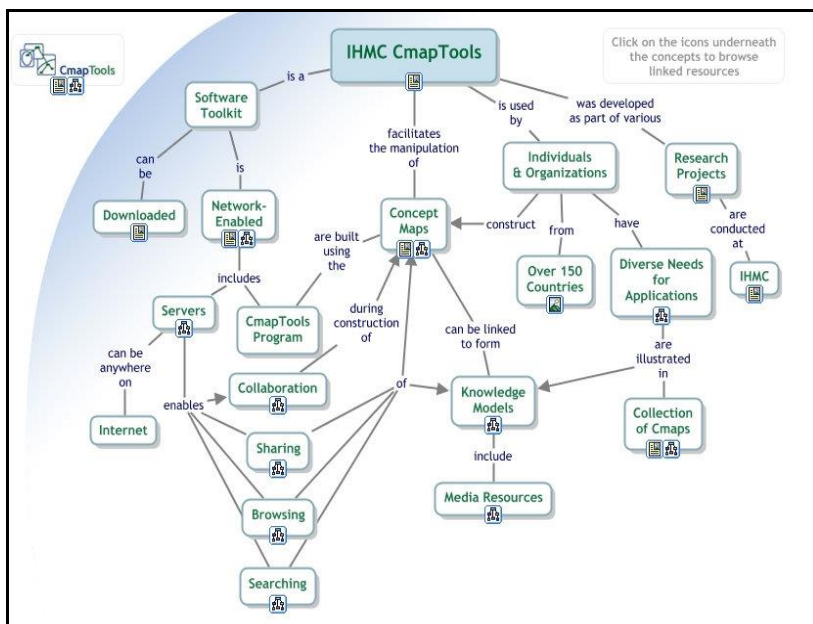
As pesquisas do IHMC estão voltadas para inteligência artificial, ciência cognitiva, modelagem e planilha de conhecimento, robótica, interfaces avançadas e displays, segurança cibernética, comunicação e colaboração, linguística, sistemas de aprendizagem mediada por computador, compreensão de dados inteligentes, agentes de *softwares*, representação do conhecimento e outras áreas afins.

A partir das pesquisas realizadas pelo IHMC, seus pesquisadores desenvolveram e continuam a implementar um interessante produto que, no contexto deste trabalho, merece uma análise mais detalhada: o *CmapTools*, que consiste em um *software* que permite construir, navegar, compartilhar e criar modelos de conhecimento representados como mapas conceituais. O *CmapTools* é distribuído gratuitamente pelo

IHMC, que o disponibiliza em conjunto com outras ferramentas com o objetivo de proporcionar ambientes colaborativos e prover meios de permitir que usuários construam e compartilhem conhecimento expresso por meio de mapas conceituais.

O site¹⁸ do IHMC *CmapTools* disponibiliza um mapa para descrever a ferramenta por eles desenvolvida e para ilustrar o funcionamento de mapas conceituais como interface visual para organização e disponibilização de informações, como pode ser visto na figura 23.

Figura 23 – Mapa Conceitual (IHMC *CmapTools*)



Fonte: http://cmaskm.ihmc.us/servlet/SBReadResourceServlet?rid=1064009710027_1637638703_27098&partName=htmltext

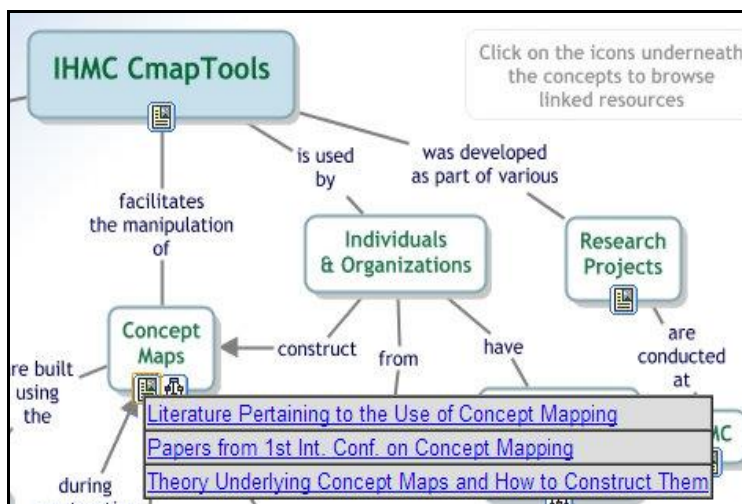
¹⁸ <http://ftp.ihmc.us/>

Na figura 23 há uma visualização baseada no modelo de interface mapa conceitual. Essa visualização é formada por conceitos, representados de forma hierárquica e por relacionamentos entre esses conceitos, indicados por uma linha que os conecta e por palavras sobre as linhas que especificam as relações entre eles. A estrutura dessa visualização auxilia na compreensão do contexto em que está inserido cada conceito e na identificação de como estão interligados entre si.

Na figura 23, o relacionamento entre conceitos busca transmitir, de forma estruturada e visual, informações referentes ao IHMC *CmapTools* como seu propósito, em que consiste, como é construído, as facilidades trazidas pelo seu uso etc.

O mapa conceitual construído pelo IHMC *CmapTools*, além de apresentar uma estratégia cognitiva para representação de informações, oferece, também, recursos para aprimorar os mapas, ou seja, permite adicionar recursos como: sons, imagens, vídeos, textos e, até mesmo, outros mapas para melhor detalhar os conceitos. Para ilustrar essa questão, tem-se a figura a seguir, que demonstra a existência de dois recursos (ícones) inerentes ao conceito *Concept Maps*. A partir de um deles, o ícone à esquerda, pode-se ter acesso às opções de documentos com conteúdos específicos, referentes a esse conceito.

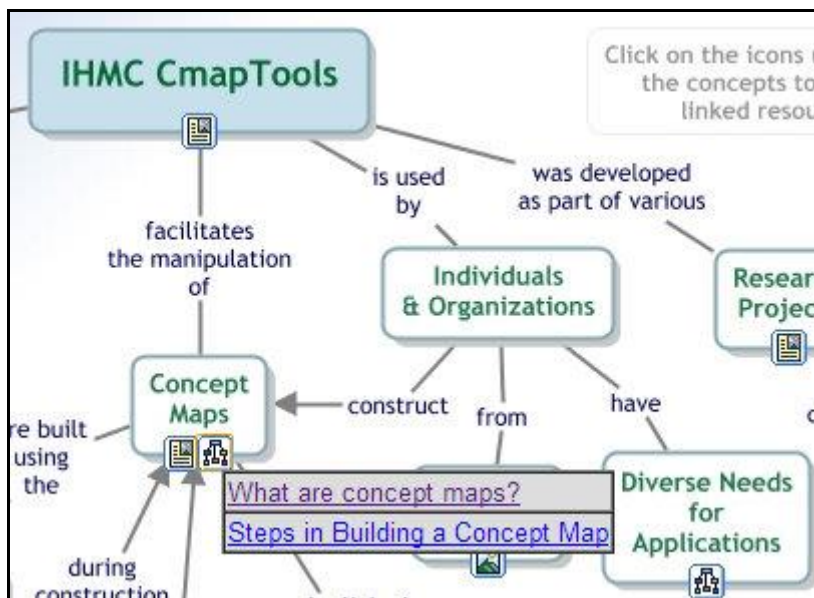
Figura 24 – Ícones para acesso a documentos no mapa conceitual



Fonte: http://cmapskm.ihmc.us/servlet/SBReadResourceServlet?rid=1064009710027_1637638703_27098&partName=htmltext

Já o ícone localizado à direita (na parte inferior do conceito *Concept Maps*), ilustrado na figura 25, direciona a navegação para novos mapas conceituais, especificamente estruturados para o conceito em questão, isto é, *Concept Maps*.

Figura 25 – Ícone para visualização de mapa conceitual



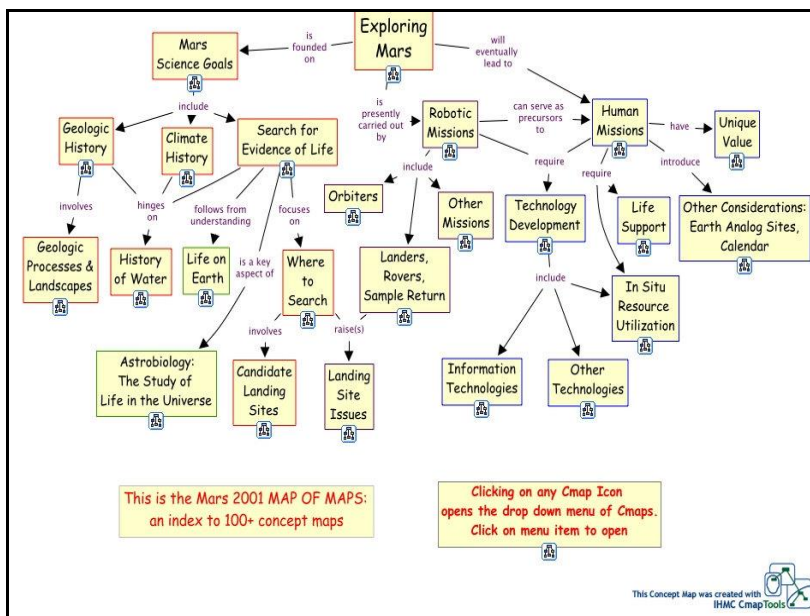
Fonte: http://cmapskm.ihmc.us/servlet/SBReadResourceServlet?rid=1064009710027_1637638703_27098&partName=htmltext

A utilização desses recursos, bem como de outros, como sons, imagens e vídeos potencializa as funções do mapa conceitual na representação visual e organizada de informações.

Nesse contexto, buscando compreender, de forma prática, como o mapa conceitual pode ser utilizado por um sistema para representação de informações, apresenta-se a seguir uma visualização, que se refere a um dos mapas conceituais adotados pelo site¹⁹ do Centro de Exploração de Marte da NASA (*CMEX Mars*).

¹⁹ <http://cmex.ihmc.us/cmex/table.html>

Figura 26 – Mapa conceitual do CMEX Mars



Fonte:

http://cmaps.nasacmex.ihmc.us/rid=1025201388710_1762691268_2257/Exploring%20Mars%20-%20Map%20of%20Maps.cmap

O ponto interessante do site do *CMEX Mars* é que ele apresenta o seu índice de assuntos por meio de um mapa de mapas, ou seja, proporciona a visualização do conteúdo disponível a partir de uma única interface, que representa um conjunto de conceitos em que cada conceito tem seu mapa específico, funcionando, assim, como um *menu* de mapas representativos, como ilustrado na figura 26.

Em síntese, os mapas conceituais, apesar de não oferecerem uma caixa de texto para busca por informações, são capazes de transmitir informações de forma organizada e

contextualizada, acarretando em uma compreensão clara do que está sendo transmitido.

6.4.1 Vantagens oferecidas pela interface visual do IHMC CmapTools

Os mapas conceituais podem ser considerados poderosas ferramentas para visualização e descobrimento de informações, pois permitem a estruturação de conceitos, ou seja, a disposição de conteúdos de maneira organizada e, conseqüentemente, possibilitam uma compreensão significativa acerca das informações representadas. Outro aspecto fundamental do mapa conceitual, como abordado anteriormente, é a possibilidade de ligação entre conceitos, que evidencia os relacionamentos existentes entre o conteúdo transmitido, sendo possível visualizar como os conceitos estão conectados com os demais e em qual contexto estão inseridos, como foi ilustrado pelas visualizações geradas pelo IHMC *CmapTools*.

Assim, de forma geral, pode-se citar as duas principais características do mapa conceitual, quais sejam, a disponibilidade de conceitos de maneira organizada e a possibilidade de ligação entre eles. Mediante essas características tem-se uma interface visual amigável, na qual o usuário pode navegar em busca de informação, bem como pode descobrir novas informações a partir do que está sendo transmitido.

Quanto à navegação a ser realizada no mapa conceitual, vale ressaltar que sua estrutura oferece ao usuário uma interface atrativa e interativa, por meio da qual a navegação pode ser realizada, proporcionando a compreensão e descoberta de informações.

Navegando pelo mapa, o usuário tem, ainda, a possibilidade de adquirir mais informações sobre determinados conceitos (a partir dos ícones que fornecem acesso aos documentos inerentes a cada conceito) e, também, de visualizar novos mapas conceituais, que remetem para a temática específica de um conceito em questão. Além dessas oportunidades, os mapas ainda comportam a adoção de imagens, filmes e sons, em contextos variados. Com tudo isso, justifica-se a utilização de mapas conceituais como ricas ferramentas na representação, organização, visualização e recuperação de informação.

6.4.2 Análise da interface visual quanto à expressividade e efetividade

Apesar de o mapa conceitual do IHMC *CmapTools* apresentar uma estrutura eficaz na representação de informações, essa interface visual pode ter sua expressividade comprometida pelo fato de disponibilizar uma grande quantidade de conceitos, o que pode acarretar em uma sobrecarga de informações, afetando a localização de conteúdos relevantes. Tendo em vista que essa interface não oferece como alternativa uma caixa de texto para a realização de buscas específicas, pode ocorrer que as informações

relevantes ao usuário venham a estar localizadas em meio a diversas outras informações, fora do contexto de interesse do usuário.

Na mesma linha de pensamento, quando se refere à efetividade do mapa conceitual do IHMC *CmapTools*, pode-se argumentar que essa visualização é efetiva quando não apresenta um número elevado de conceitos e relações, pois o grande volume de relações e conceitos dispostos pode dificultar a compreensão das informações, não sendo de fácil entendimento para os usuários.

6.4.3 Análise dos elementos de OI necessários para o emprego de interface visual baseada no mapa conceitual

Para a construção de uma interface visual baseada em mapa conceitual, é importante considerar que o conteúdo a ser transmitido necessita estar organizado e estruturado adequadamente, para que possa ser representado e visualizado em forma de mapa. Para isso, torna-se pertinente a adoção de um instrumento de OI que possibilite a organização de conceitos e, principalmente, que permita estabelecer as ligações existentes entre eles.

Levando em conta que o mapa conceitual é formado por uma estrutura complexa de conceitos dispostos em categorias, representando desde os conceitos mais gerais até os mais específicos e que, além disso, o mapa conceitual é caracterizado pelos relacionamentos estabelecidos entre os conceitos, resultando em uma estrutura rica, porém

complexa, se faz necessário, então, a adoção de um instrumento que ofereça o suporte necessário a esse contexto. Sob essa ótica, pode-se citar como alternativa de instrumento as ontologias, já que são conhecidas pela potencialidade em organizar e representar informação.

As ontologias, como menciona Silva (2008), são estruturas que congregam conjuntos de conceitos padronizados, termos e definições, evidenciando a forma com que esses conceitos estão relacionados entre si. Nessa linha de pensamento, Brandão e Lucena (2002) enfatizam que o uso de ontologias possibilita não só a definição de conceitos num domínio de conhecimento, como também permite a definição de relacionamentos entre esses conceitos e de inferências básicas os envolvendo.

Considerando esse âmbito, têm-se as ontologias como importantes aliadas na organização, categorização e facilitação do estabelecimento de relações entre conceitos (propósito central do mapa conceitual), acarretando, assim, em melhorias tanto na representação como na recuperação de informação.

6.5 WEBSOM - método para construção de mapas visuais

WEBSOM é um método de construção de mapas visuais para representação de documentos textuais. O WEBSOM organiza coleções de documentos e permite a exploração e recuperação desses documentos por meio de uma interface visual baseada na técnica cartografia temática.

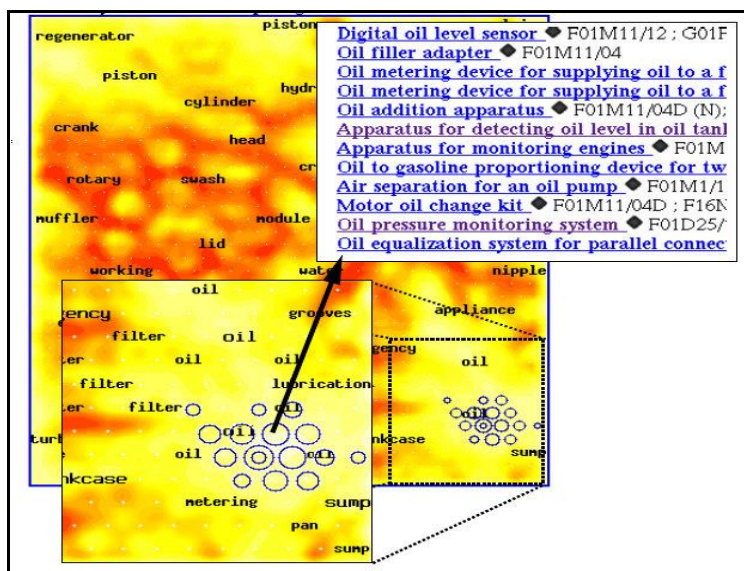
O modelo de interface de VI cartografia temática é utilizado, frequentemente, por sistemas de informação, na elaboração de mapas temáticos sobre determinados fenômenos, geralmente geográficos, com o intuito de representar informações de maneira visual. Em contrapartida, a WEBSOM tem como propósito organizar e representar documentos textuais, proporcionando a visualização e exploração de conjuntos de documentos e facilitando a sua localização e recuperação, razão pela qual se destaca, aqui, a WEBSOM, para uma análise mais aprofundada acerca da sua estrutura e funcionalidade.

No método WEBSOM, os documentos armazenados na base de dados do sistema são representados e organizados por meio de termos para produzir um mapa de documentos. Esses termos indicam o contexto dos documentos e ficam agrupados em pontos do mapa de acordo com a similaridade entre os conteúdos que representam, ou seja, a proximidade entre os termos indica que eles abordam temáticas relacionadas. Essa relação de associação entre os documentos é medida pela coocorrência entre termos. Dessa forma, os documentos semelhantes estão posicionados em grupos próximos uns dos outros no mapa.

A figura 27 ilustra o mapa gerado pelo WEBSOM para fins de recuperação por navegação, este mapa é um exemplo de cartografia temática que retrata o espaço de documentos, isto é, mostra a relação de similaridade entre documentos e revela temas principais tratados por grupos de documentos.

Ao clicar sobre uma região do mapa, é possível obter um zoom dessa área, o que leva a uma visualização mais aprofundada acerca do grupo de documentos pertinentes à temática da região selecionada, permitindo assim a exploração de vertentes mais específicas. A figura 28, a seguir, ilustra essa ideia.

Figura 28 – Zoom sobre uma área específica do mapa



Fonte: Lagus et al. (1999)

Também pode ser visto por meio da figura 28 que, para ter acesso a um documento representado em um determinado grupo, basta clicar no termo disposto no nível de menor zoom do mapa; a partir disso, são listados os documentos e o seu conteúdo pode ser consultado.

6.5.1 Vantagens oferecidas pela interface visual do WEBSOM

A visualização construída pelo método WEBSOM, por meio de sua estrutura em forma de mapa, baseada na cartografia temática, apresenta vantagens para os usuários que dela fazem uso, uma vez que proporciona uma visão geral do conjunto de conteúdos disponíveis, bem como permite identificar na visualização os conteúdos semelhantes ou os que abordam temáticas relacionadas.

Assim, uma das vantagens trazidas pelo mapa do WEBSOM é que ele fornece uma visão completa das temáticas contidas nos documentos armazenados, permitindo que o usuário navegue por entre os assuntos, na busca por informações que sejam de seu interesse. Desse modo, ao acessar um sistema que faça uso dessa interface visual, pode-se ter uma noção dos assuntos abordados pelo sistema e, também, de como estão estruturados.

Além da possibilidade de visualizar o conteúdo disponível de uma forma geral, a interface do WEBSOM proporciona a identificação dos assuntos com temáticas semelhantes, ou melhor, por meio da posição ocupada pelos grupos de documentos no mapa, tem-se um ponto de partida para explorar documentos relacionados. Assim, por meio da proximidade entre os termos que representam os documentos no mapa, o usuário consegue localizar e descobrir documentos relacionados à sua busca inicial.

É importante enfatizar, ainda, que o WEBSOM é dotado de uma rica estrutura, que oferece a base para a pesquisa e exploração de conteúdos específicos, pois o

usuário tem a oportunidade, por meio do zoom, de visualizar mais detalhes acerca do espaço de documentos subjacentes. Os termos agrupam os documentos dispostos na interface do WEBSOM por temáticas mais gerais, no entanto, tem-se a possibilidade de acessar e explorar as temáticas mais específicas até o menor nível do zoom, onde se obtêm os documentos completos sobre as determinadas regiões.

Dessa forma, a interface baseada na cartografia temática disponibiliza uma navegação interativa, por meio da qual se pode localizar e recuperar conteúdos pertinentes, mediante uma estrutura gráfica.

6.5.2 Análise da interface visual quanto à expressividade e efetividade

Analisando a interface visual do WEBSOM quanto aos atributos capazes de caracterizar o seu bom desempenho - expressividade e efetividade - entende-se, em primeiro lugar, que esta interface não pode ser considerada expressiva, tendo em vista que uma visualização é expressiva quando é capaz de apresentar todas as informações de interesse do usuário e nenhum dado a mais ou a menos fora do contexto. Nesse caso, a interface do WEBSOM não cumpre esse propósito, já que, ao clicar em um determinado termo não se tem todos os documentos a ele relacionados, mas somente os documentos localizados na região selecionada, dentro do contexto ali representado.

Por outro lado, a interface do WEBSOM pode ser considerada efetiva, uma vez que permite ao usuário conhecer os assuntos tratados e encontrar documentos relacionados. No entanto, essa efetividade pode ser comprometida quando se tem uma grande quantidade de termos expressos no mapa, o que pode acarretar em uma sobrecarga de informação, dificultando, assim, a compreensão das informações transmitidas pela estrutura e a localização de conteúdos desejados.

6.5.3 Análise dos elementos de OI necessários para o emprego de interface visual baseada em cartografia temática

Para que uma interface baseada em cartografia temática possa ser construída, se faz necessário que os documentos a serem representados de forma visual sejam previamente organizados, o que consiste em um passo crucial para garantir uma interface expressiva que possibilite a exploração e o descobrimento de informações úteis.

Sendo assim, é importante compreender, em primeiro lugar, que os documentos na estrutura da cartografia temática são representados por termos e esses termos são dispostos de acordo com a proximidade temática entre os assuntos dos documentos por eles representados, buscando evidenciar aqueles que compartilham conteúdos afins. É relevante esclarecer, também, que os documentos são agrupados com base na coocorrência estatística como indicador de proximidade ou similaridade entre eles. Assim, os termos

considerados semelhantes tematicamente são agrupados em blocos, numa estrutura de mapa visual, onde são dispostos como pontos num espaço cartesiano.

Após compreensão da estrutura da cartografia temática, torna-se clara a necessidade de adoção de um instrumento que permita representar os documentos por meio de um vocabulário controlado e que seja capaz de expressar as relações associativas entre os documentos. Sob essa perspectiva, pode-se citar como instrumento adequado ao contexto da cartografia temática as ontologias, tendo em vista que, como já mencionado por Silva (2008), as ontologias são estruturas que congregam conjuntos de conceitos padronizados, termos e definições aceitas por uma comunidade particular e a forma como esses conceitos estão relacionados entre si.

Assim, a partir de uma representação adequada dos documentos, mediante adoção de ontologia e do estabelecimento de temáticas relacionadas, o sistema poderá estabelecer critérios para identificar a coocorrência entre termos, ou seja, para identificar os documentos nos quais mais ocorrem termos associados em comum. A partir disso, os documentos podem, então, ser agrupados e representados em forma de mapa visual, evidenciando os temas principais tratados por esses grupos de documentos, bem como indicando a proximidade temática entre eles.

6.6 Sistema de busca da BRAPCI

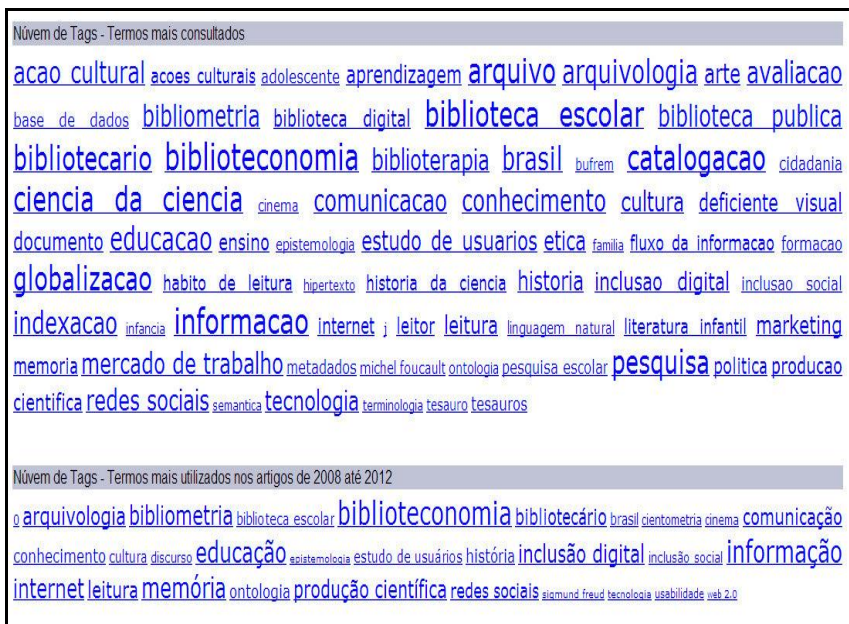
A Base de Dados Referenciais de Artigos de Periódicos em Ciência da Informação (BRAPCI) é um produto de informação resultante do projeto de pesquisa intitulado “Opções metodológicas em pesquisa: a contribuição da área da informação para a produção de saberes no ensino superior”, que tem como objetivo subsidiar estudos e propostas de construção de conhecimentos na área de CI.

O propósito da BRAPCI é oferecer à comunidade acadêmica e científica o acesso a artigos de periódicos da área de CI e, para isso, identifica títulos de periódicos da área e indexa seus artigos. Dessa forma, essa base de dados é constituída, atualmente, por cerca de 8.292 textos publicados em 37 periódicos nacionais, impressos e eletrônicos, sendo 28 desses periódicos ativos e 9 históricos, ou seja, inativos.

O diferencial da BRAPCI é que, além de disponibilizar uma caixa de texto, por meio da qual pesquisas podem ser realizadas, essa base adota, também, visualizações em forma de nuvem de *tags* como uma alternativa amigável para a busca e identificação de conteúdos.

Para a realização de buscas por meio de interface visual, a página inicial do site²⁰ da BRAPCI oferece duas nuvens de *tags*, como pode ser visto na figura 29.

²⁰ <http://www.brapci.ufpr.br/index.php>

Figura 29 – Nuvens de Tags do BRAPCI

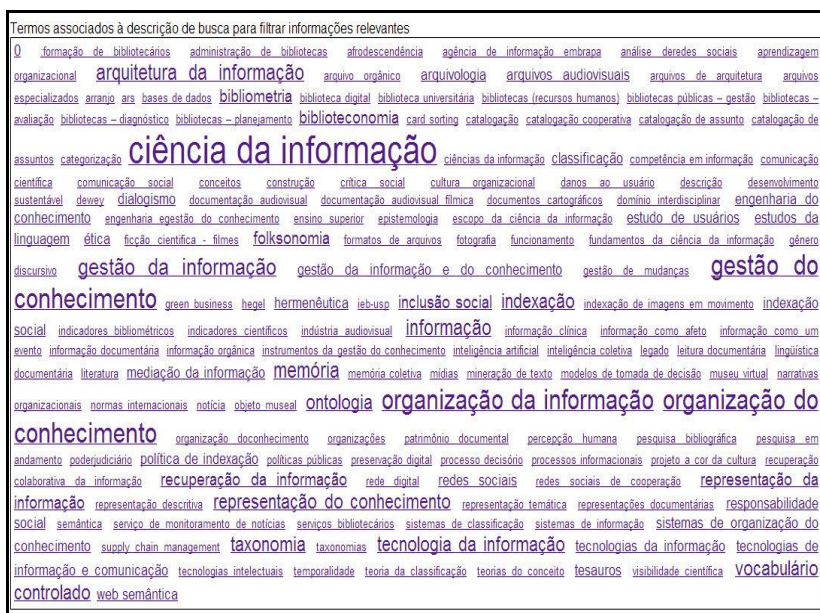
Fonte: <http://www.brapci.ufpr.br/index.php>

A primeira nuvem apresenta as palavras (*tags*) mais consultadas no momento, caracterizadas como as *tags* mais populares; a segunda nuvem apresenta as *tags* estabelecidas pelo sistema da base como sendo as mais utilizadas nos artigos do ano de 2008 até o ano de 2012. Por meio dessas nuvens pode-se ter uma direção para algum tema específico antes de submeter uma busca à base de dados.

Além dessa opção, tem-se a possibilidade de realizar uma busca submetendo palavras-chave por meio de uma caixa de texto; dessa forma, é gerada uma nuvem com *tags* referentes à busca submetida. A nuvem de *tags* gerada a

partir de uma consulta específica apresenta temáticas relacionadas ao que foi pesquisado, indicando novos caminhos a serem percorridos ou alternativas relacionadas com a pesquisa em questão. Como exemplo, a figura 30, a seguir, ilustra a nuvem de *tags* gerada para uma busca realizada por meio do termo “Organização da Informação”, nela, pode-se visualizar algumas *tags* associadas ao termo pesquisado, como “ciência da informação”, “representação do conhecimento”, “gestão da informação”, entre outros.

Figura 30 –Nuvem de *tags* gerada a partir de busca por meio do termo organização da Informação



Fonte: http://www.brapci.ufpr.br/search_result.php

Outro aspecto interessante a ser analisado na nuvem de *tags* da figura 30, diz respeito as que *tags* que possuem um destaque maior quanto ao tamanho da fonte. Esse aspecto é explicado pelo próprio desenvolver da BRAPCI ao esclarecer que, no caso específico dessa base, as *tags* que possuem uma fonte em tamanho maior indicam maior incidência de palavras-chave sobre o tema representado pela *tag*, ou seja, uma maior incidência de artigos sobre a temática representada (informação verbal)²¹.

Esclarecido esse aspecto, vale mencionar que para ter acesso aos artigos indexados pela BRAPCI pertinentes às *tags* dispostas na nuvem, é necessário, apenas, selecionar a *tag* de maior interesse e, a partir disso, são listados os artigos referentes a ela. Essa listagem apresenta os artigos por meio do título, resumo e palavras-chave, disponibilizando a opção de download dos artigos em PDF. Além disso, a BRAPCI oferece meios para obtenção de referência bibliográfica de acordo com as normas da ABNT e VANCOUVER.

6.6.1 Vantagens oferecidas pela interface visual da BRAPCI

Mediante a adoção da nuvem de *tags*, como um modelo de interface visual, a BRAPCI disponibiliza ao seu usuário uma alternativa dinâmica e amigável, por meio da qual se pode visualizar, descobrir e recuperar informação.

²¹ Informação fornecida por Rene Fausyino Gabriel Junior, em Recife, em fevereiro de 2014.

As nuvens apresentadas na página inicial da BRAPCI, onde estão dispostas as *tags* mais consultadas e utilizadas no momento, têm o propósito de informar o usuário acerca das temáticas mais populares, mas também podem fornecer pistas, sugerindo termos para possíveis buscas. Assim, caso o usuário se dirija à base de dados em busca de informação, mas não tenha uma consulta formulada, ou não tenha uma direção clara do caminho a ser percorrido para alcançar seu objetivo, a nuvem de *tags* pode auxiliá-lo, oferecendo sugestões de pesquisa para a localização de informações úteis.

No entanto, essas *tags* dispostas na página inicial da BRAPCI não representam todas as temáticas disponíveis, mas as mais consultadas de uma forma geral, como mencionado anteriormente. Diante disso, pode-se submeter uma consulta por meio da caixa de texto, para que uma busca por temáticas específicas seja realizada e, então, é gerada uma nova nuvem de *tags* com foco no termo pesquisado. Essa nuvem oferece uma visão ampla do conjunto de *tags* mais acessados acerca do que foi pesquisado e, ao mesmo tempo, permite identificar as especificidades dentro do domínio em questão.

Nesse contexto, a interface em forma de nuvem de *tags* é capaz de orientar o usuário para temáticas específicas, auxiliando-o a filtrar conteúdos relevantes. Pode-se citar, como exemplo, a realização de uma busca por meio do termo “Representação da Informação”, que vai resultar em *tags* como “classificação”, “indexação”, “catalogação”, “representação temática”, “linguagem documentária”, entre

outras. Assim, mesmo que o usuário tenha uma ideia superficial do que deseja buscar, por meio das *tags* dispostas na nuvem ele poderá especificar ou focar em uma linha de seu interesse.

Por outro lado, é possível afirmar que a nuvem de *tags*, por oferecer uma visão ampla dos assuntos disponíveis, possibilita, também, a visualização de termos associados à descrição de busca, o que a torna uma poderosa aliada para o usuário, já que o informa acerca das temáticas semelhantes à sua questão inicial. Seguindo o mesmo exemplo anteriormente citado, para uma busca realizada por meio do termo “Representação da Informação”, obtém-se termos a ele relacionados como “organização do conhecimento” e “recuperação da informação”. Sob essa perspectiva, a nuvem de *tags* enriquece a busca e facilita a descoberta de novas vertentes que poderiam não ser do conhecimento do usuário, mas que podem vir a ser pertinentes à sua necessidade.

6.6.2 Análise da interface visual quanto à expressividade e efetividade

Para análise da interface visual da BRAPCI quanto a sua expressividade e especificidade é preciso ressaltar, em primeiro lugar, que o objetivo dessa interface, em forma de nuvem de *tags*, não é o de apresentar todos os termos referentes ao conteúdo existente na base de dados, mas os mais discutidos ou acessados. Sob essa ótica, a interface da BRAPCI pode ser considerada expressiva para os usuários

que desejam explorar a coleção de documentos por meio dos assuntos mais tratados ou mais procurados e seus respectivos artigos.

Quanto à efetividade da interface visual da BRAPCI, pode-se argumentar que se trata de uma interface efetiva, uma vez que organiza as *tags* de maneira alfabética, estruturando-as claramente e, conseqüentemente, proporcionando uma fácil navegação, por meio da qual a identificação e a recuperação de documentos pode ser realizada sem erros de interpretação.

6.6.3 Análise dos elementos de OI necessários para o emprego de interface visual baseada na nuvem de tags

Para análise da nuvem de *tags* quanto à organização da informação adequada à sua funcionalidade, é importante considerar que esse modelo de interface não tem como propósito estabelecer relações entre as *tags* (palavras) dispostas na nuvem, ou seja, as *tags* não possuem entre si relacionamentos por hierarquias, associações ou de qualquer tipo. As relações associativas que podem ocorrer são aquelas estabelecidas entre os termos pesquisados em uma busca e as *tags* dispostas na nuvem, que representam temáticas relacionadas ou semelhantes aos termos pesquisados.

Esclarecido esse fato e retomando a discussão acerca da OI no contexto da nuvem de *tags*, entende-se como indicado que o sistema responsável por adotar esse tipo de

interface faça uso de um instrumento de OI que sirva de base para organização das *tags* a serem dispostas de forma visual. Nesse âmbito, têm-se os cabeçalhos de assunto, como instrumentos que uniformizam a descrição sintética dos assuntos contidos nos documentos, servindo como pontos de acesso para sua localização.

Os cabeçalhos de assunto são elaborados a partir de palavras simples ou compostas (ordenadas em lista alfabética), geralmente extraídas do próprio documento, para descrever seus assuntos. Nas listas formadas pelos cabeçalhos de assunto podem ser encontradas, ainda, remissivas de orientação para assuntos relacionados como, por exemplo: “linguagem documental”, ver também “classificação”, “documentação”, “linguística” (GUINCHAT; MENOU, 1994).

As listas de cabeçalhos de assunto podem ser, portanto, úteis à organização da informação na nuvem de *tags*, visto que podem auxiliar na estruturação, representação e padronização das *tags*, bem como possibilitar a identificação e recuperação de documentos sobre uma determinada temática ou sobre temáticas relacionadas, pois apesar dos termos na lista de cabeçalho de assunto serem independentes entre si, as remissivas existentes permitem orientar o usuário para noções próximas do assunto que busca, como afirmam Guinchat e Menou (1994). Assim, ainda segundo os autores, tem-se a possibilidade de recuperar um documento cujo assunto se conhece e saber que documentos a unidade de informação possui sobre determinado assunto.

6.7 Relação entre OI e VI nos sites de SRIs que fazem uso de interface visual

Com base nas análises realizadas anteriormente acerca dos SRIs que adotam interfaces visuais em sua estrutura, é possível argumentar que, embora os sites de SRIs analisados apresentem propostas e funcionalidades diversificadas, todos eles caminham para os mesmos objetivos – facilitar a visualização, navegação, descoberta e recuperação de informações. Para alcançar esses objetivos, os sites adotam técnicas e modelos de interface visual que melhor se adéquam aos seus contextos e particularidades. Em decorrência da adoção das interfaces visuais, esses sites oferecem diversos benefícios aos usuários, como por exemplo:


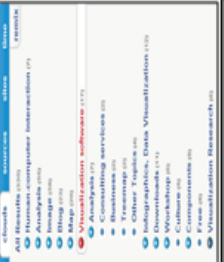
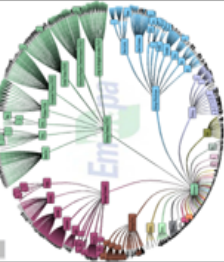
- Oferecem uma maneira amigável e dinâmica de interação com as informações dispostas;
- Auxiliam na compreensão de como estão estruturadas as informações;
- Facilitam na localização de informações precisas;
- Contribuem para descoberta e dedução de novas informações,
- Permitem a visualização de conteúdos diferentes que antes eram desconhecidos, mas que podem vir a ser pertinentes à necessidade do usuário;
- Minimizam a sobrecarga de informação sobre o usuário.


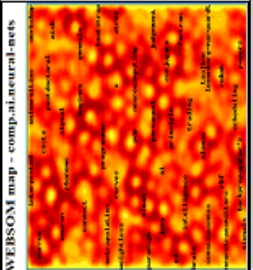

Outro ponto importante a ser analisado quando se refere à aplicação de interfaces de VI em SRIs diz respeito aos aspectos que envolvem a organização prévia das informações a serem representadas.

A esse respeito, identificou-se a influência que a OI exerce, por meio de seus instrumentos, para auxiliar na organização e representação de conteúdos que serão transmitidos em estruturas visuais, confirmando, dessa forma, que sistemas que fazem uso de visualizações para transmitir informação requerem métodos de OI adequados para dar suporte às representações visuais de informação e para permitir a recuperação de conteúdos, já que, como afirma Zhang (2008), os métodos de organização e armazenamento de informações afetam e determinam os métodos de recuperação de informação.

Nesse contexto, buscando demonstrar, de maneira resumida, os resultados obtidos por meio da análise anteriormente realizada nos sites de SRI quanto à interação entre OI e VI, apresenta-se, a seguir, um quadro que permite visualizar as técnicas e modelos de interface visual adotados por cada um dos sites, bem como os instrumentos de OI adequados para a construção e funcionamento de visualizações expressivas.

Quadro 2 – Relação entre OI e VI na aplicação de interface visual

Site de SRI	Técnica ou modelo de VI	Interface Visual	Instrumentos de OI para o emprego de visualização
AquaBrowser da biblioteca da Universidade de Edimburgo	Desenho de Grafos		Adoção de um tesouro ou taxonomia como instrumento de OI que permita uma representação e organização adequada de termos e que estabeleçam as relações existentes entre conceitos.
Mecanismo de busca YIPY	Desenho de Árvores		Adoção de uma taxonomia como instrumento de OI que permita a categorização de conteúdos semelhantes, evidenciando as relações hierárquicas existentes entre eles.
Infoteca-e da Embrapa	Browser Hiperbólico		Adoção de uma taxonomia como instrumento capaz de organizar termos em categorias da mais genérica a mais específica, estabelecendo relações hierárquicas entre eles.

Website do IHMC Camp Tools	Mapa Conceitual		Adoção de ontologia como instrumento que possibilite congregar conjuntos de conceitos padronizados, termos e definições, evidenciando a forma com que esses conceitos estão relacionados entre si.
WEBSOM - método para construção de mapas visuais.	Cartografia Temática		Adoção da ontologia enquanto instrumento que permite a representação e a categorização de documentos, bem como o estabelecimento de temáticas relacionadas; adoção de critérios para identificar a coocorrência entre termos.
Sistema de busca da BRAPCI	Nuvem de Tags		Adoção de listas de cabeçalhos de assunto para auxiliar na organização, estruturação e representação das tags que servirão como pontos de acesso para localização de documentos.

Fonte: elaborado pela autora

Assim, por meio do quadro 2, é possível identificar quais instrumentos de OI podem ser adequados para auxiliar os sistemas que adotam técnicas ou modelos de interface visual específicos. Tais instrumentos foram indicados por possuírem capacidades de organização que se adéquam à estrutura da visualização adotada pelo sistema, isto é, à necessidade requerida pela técnica ou modelo de interface em questão, como por exemplo: categorização de conteúdos semelhantes; estabelecimento de relações hierárquicas; congregação de conjuntos de conceitos padronizados, termos e definições; estabelecimento de relações associativas.

Sob essa perspectiva, é importante elucidar que a atuação desses instrumentos de OI é direcionada para a organização das informações a serem transmitidas e que essa organização poderá ser refletida de maneira visual, podendo acarretar em visualizações expressivas, capazes de permitir representação facilitada, por meio de estrutura simples e recuperação satisfatória de informações.

7 CONSIDERAÇÕES FINAIS

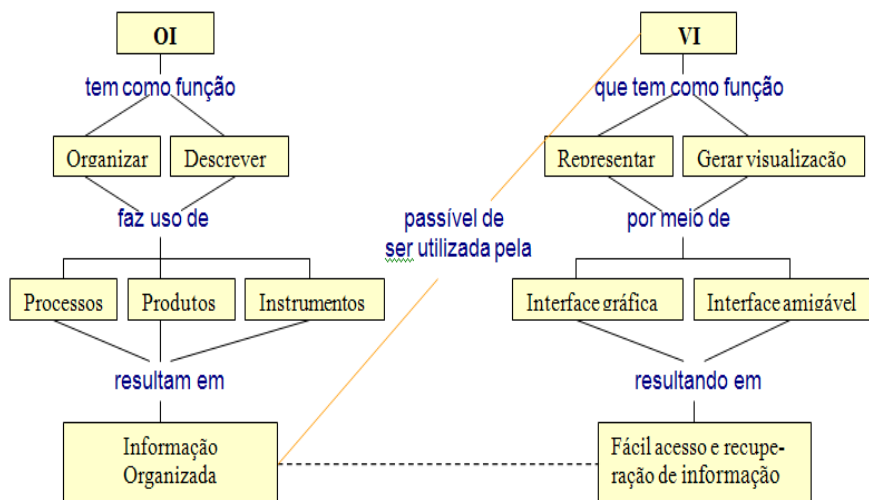
Os SRIs, enquanto sistemas que lidam com informação registrada e que trabalham com o propósito de permitir, a partir de um universo de documentos disponíveis, a recuperação daqueles que venham a ser úteis para satisfação de uma necessidade informacional, exercem um importante papel em uma sociedade caracterizada pela crescente produção e disseminação de informação, amplamente influenciada pelas tecnologias de informação e comunicação disponíveis.

Nesse contexto, os SRIs podem ser vistos como sistemas de armazenamento, preservação e acesso à memória humana produzida, registrada e por eles armazenada, o que confere a eles, como visto anteriormente, o título de sistemas detentores de memória. Esse título atribui a eles uma importante missão, a de tornar acessível uma quantidade cada vez maior de conteúdo informacional, o que implica numa necessidade por meios precisos de acesso à informação – elemento vital às atividades humanas – bem como de maior controle sobre essa informação, com vista a prover a sua socialização e recuperação.

Considerando essas necessidades, partiu-se do pressuposto que a OI e a VI, enquanto áreas que assumem uma relação direta com a recuperação da informação, podem ser consideradas poderosas aliadas para os SRIs. Neste sentido, buscou-se uma análise mais aprofundada acerca dessas áreas, o que resultou em uma compreensão mais clara sobre as particularidades de cada uma e tornou possível

identificar a potencial ligação entre elas. A figura 31 permite, por meio de uma visualização sucinta, ilustrar essa questão.

Figura 31 - Análise do foco das áreas de OI e VI



Fonte: elaborado pela autora

A partir do que foi abordado na revisão de literatura acerca da OI e VI, foi possível evidenciar (conforme demonstrado na figura 31) que a OI, enquanto área que se preocupa com a organização, descrição e tratamento da informação, fazendo uso, para isso, de processos, produtos e instrumentos, pode subsidiar a construção de interfaces de VI, cujo objetivo é representar informações de forma gráfica, facilitando o acesso e recuperação de conteúdos informacionais, por meio de uma interface amigável.

Dessa forma, entende-se que a OI é capaz de oferecer um suporte para que a VI exerça sua função de forma satisfatória, pois para que a VI potencialize o acesso e a recuperação da informação, o ideal é que essa informação esteja prévia e adequadamente organizada. Entende-se, assim, que a OI e a VI possuem focos específicos, mas que podem caminhar em harmonia para um objetivo comum: possibilitar o acesso e recuperação eficaz de informação organizada.

Foi exatamente nesse contexto que se desenvolveu o objetivo proposto por esta pesquisa, ou seja, buscou-se analisar a OI e a VI sob o ponto de vista da contribuição dessas áreas para os SRIs. Para isso, como validação do que foi pesquisado durante a revisão de literatura, foram analisados sites de SRIs que fazem uso de interfaces de VI, com o intuito de evidenciar as vantagens oferecidas pelo uso de visualizações nesses sistemas, bem como permitir a análise dos requisitos de OI necessários para tornar possível a construção das interfaces visuais e a recuperação de conteúdos informacionais relevantes.

Por meio da análise desses sites pode-se comprovar que, de fato, tem-se nas interfaces de VI ricas alternativas para representação de conteúdo informacional, disponibilizando uma maneira interativa para visualização, localização e recuperação das informações armazenadas. Além disso, sob a perspectiva da OI, foi possível compreender a influência que a adoção de seus instrumentos exercem na organização e representação de conteúdos que serão transmitidos em estruturas visuais. Verificou-se que a adoção

de instrumentos adequados pode resultar em uma recuperação eficaz de informações, corroborando, assim, que os recursos de OI afetam diretamente a recuperação da informação, conforme foi ressaltado ao longo da pesquisa.

Neste momento, é importante destacar que, embora não se possa afirmar que os sites analisados adotam em sua estrutura os instrumentos verificados como adequados nesta pesquisa, tais instrumentos foram sugeridos, pois entende-se que por meio da adoção deles se tem uma maior precisão na representação da informação, o que poderá possibilitar uma maneira expressiva de transmitir e recuperar informações por meio de interface visual, minimizando as chances de conteúdos relevantes permanecerem ocultos.

Em síntese, tem-se que os SRIs, enquanto sistemas que detêm o importante papel de guardiões e propagadores da memória, no que se refere às informações registradas, necessitam de uma maior atenção à maneira como são abordados o acesso e a recuperação de informação. Essas necessidades estão alinhadas à questão básica da CI, que gira em torno do acesso à informação, ou facilitação desse acesso. Assim, fundamentando-se no fato de que a CI está envolvida e empenhada nos aspectos que envolvem a recuperação da informação, bem como os meios de processá-la para otimizar sua acessibilidade e uso, considera-se, nesse cenário, a OI e a VI como potenciais auxiliadoras para os SRIs, considerando os preceitos de ambas as áreas e o papel que elas exercem quando se refere à informação, sua organização, visualização e recuperação.

Fundamentando-se na pesquisa desenvolvida, bem como nos autores referenciados neste trabalho chegou-se, por fim, à compreensão de que as referidas áreas, a saber, OI e VI, convergem entre si, ou seja, partilham de objetivos comuns e, a partir disso, defende-se que essas áreas podem alcançar promissores resultados, se atuarem em parceria para otimização na recuperação da memória armazenada em SRIs.

REFERÊNCIAS

ALVARENGA, L. Representação do conhecimento na perspectiva da ciência da informação em tempo e espaço digitais. **Encontros Bibli: revista eletrônica de Biblioteconomia e Ciência da Informação**, Florianópolis, v. 8, n. 15, p. 18-40, 1º sem. 2003.

ALVES et al. Ciência da Informação, Ciência da Computação e Recuperação da Informação: algumas considerações sobre os métodos e tecnologias da informação utilizados ao longo do tempo. **Revista Eletrônica Informação e Cognição**, São Paulo, v.6, n.1, p.28-40, 2007.

ANGLO-AMERICAN CATALOGING RULES. **About AACR2**. Canada, 2006. Disponível em: <
<http://www.aacr2.org/about.html>>. Acesso em: 17 abr. 2013.

ANZOLIN, H. H. Atualizações em AACR2. In: ENCONTRO NACIONAL DOS USUÁRIOS DA REDE PERGAMUM, 9., 2007, Curitiba. **Anais...** Curitiba: PUCPR, 2007. 1 CD-ROM.

ARAÚJO JUNIOR, R. H. **Precisão no processo de busca e Recuperação da informação**. Brasília: Thesaurus, 2007.

ARAÚJO, V. M. R. H. de. Sistemas de informação: nova abordagem teórico-conceitual. **Ciência da Informação**, Brasília, v. 24, n. 1, p. 1-39, jan./abr. 1995.

ARELLANO, F. F. M. Las RDA (Resource Description and Access) y su impacto en la catalogación y los catálogos: necesidades de investigación. In: ENCUENTRO DE CATALOGACIÓN Y METADADOS, 4., 2009, México. **Anais...** México, DF: UNAM/CUIB, 2010. p. 65-84. Disponível em: <http://132.248.242.3/~publica/archivos/libros/iv_encuentro_catalogacion.pdf>. Acesso em: 18 abr. 2013.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS.
NBR6028: Informação e documentação: resumo: apresentação. Rio de Janeiro, 2003. 2p.

BAEZA-YATES, R.; RIBEIRO-NETO, B. **Modern Information Retrieval**. New York: ACM Press, 1999.

BRANDÃO, A. A. F., LUCENA, C. J. P. de. **Uma Introdução à Engenharia de Ontologias no contexto da Web Semântica**. Rio de Janeiro. Departamento de Informática, PUC-Rio, 2002. Disponível em: < ftp://ftp.inf.puc-rio.br/pub/docs/techreports/02_29_brandao.pdf >. Acesso em: 10 out. 2013.

BRASCHER, M.; CAFÉ, L. Organização da Informação ou Organização do Conhecimento?. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 9., 2008, São Paulo. **Anais...** São Paulo: ENANCIB, 2008. p. 1-14.

BRASCHER, M.; CARLAN, E. Sistemas de organização do conhecimento: antigas e novas linguagens. In: ROBREDO, J.;

BRÄSCHER, M. (Org.). **Passeios pelo bosque da informação**: estudos sobre a representação e organização da informação e do conhecimento – eroic. Brasília DF: IBICT, 2010. p. 148-176.

BRAZ, M. I. **Dispositivos de memória para informação jurídica**: análise de procedimentos de indexação. 2013. 100f. Dissertação (Mestrado em Ciência da informação) - Departamento de Ciência da Informação, Universidade Federal de Pernambuco, Recife, 2013.

CAFÉ, L.; SALES, R. Organização da informação: conceitos básicos e breve fundamentação teórica. In: ROBREDO, J.; BRÄSCHER, M. (Org.). **Passeios pelo bosque da informação**: estudos sobre a representação e organização da informação e do conhecimento – eroic. Brasília DF: IBICT, 2010. p. 115-129.

CAÑAS, A. J. et al. Concept Maps: integrating Knowledge and Information Visualization. In: Tergan, S. O.; Keller, T. (Ed.). **Knowledge and Information Visualization**: searching for Synergies. Heidelberg, NY: Springer, 2005. p. 205-219.

CARDOSO, O. N. P. Recuperação da informação. **Infocomp**, Lavras, v. 2, n. 1, p. 27-32, 2000. Disponível em: <<http://www.dcc.ufla.br/infocomp/artigos/v2.1/art07.pdf>>. Acesso em: 11 mar. 2012.

CARLAN, E.; MEDEIROS, M. B. B. Sistemas de Organização do Conhecimento na visão da Ciência da Informação. **RICI**: revista Ibero-americana de Ciência da Informação, Brasília, v. 4, n. 2, p. 53-73, ago./dez. 2011.

CASTRO, F. F. de; SANTOS, P. L. V. A. da C. Catalogação e Metadados: interlocuções nos ambientes informacionais digitais. In: ENCUESTRO DE CATALOGACIÓN Y METADADOS, 4., 2009, México. **Anais...** México, DF: UNAM/CUIB, 2010. p. 301-318.

CORRÊA, R. M. R. **Catalogação descritiva no século XXI**: um estudo sobre o RDA. 2008. 75 f. Dissertação (mestrado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2008.

DAHLBERG, I. Teoria do conceito. **Ciência da Informação**, Brasília, v.7, n.2, p. 101-107, jul./ dez. 1978.

D'ANDRÉA, C. Estratégias de produção e organização de informações na web: conceitos para a análise de documentos na internet. **Ciência da Informação**, Brasília, v. 35, n. 3, p.39-44, 2006.

DIAS, E. W. Contexto Digital e Tratamento da Informação. **DataGramZero**: revista de Ciência da Informação, Rio de Janeiro, v.2, n.5, p. 1-9, out. 2001.

DIAS, M. P.; CARVALHO, J. O. F. de. A Visualização da Informação e a sua contribuição para a Ciência da Informação. **DataGramaZero**: revista de Ciência da Informação, Rio de Janeiro, v.8, n.5, p. 1-16, out. 2007.

ELIEL, O. **Cartografia temática de artigos de periódicos nacionais da Ciência da Informação (período de 1986 a 2005)**: contribuições ao campo teórico da área. 2007.120f. Dissertação (Mestrado em Ciência da informação) – Centro de Ciências Sociais Aplicadas, Pontifícia universidade Católica de Campinas, Campinas, 2007. Disponível em:<http://www.bibliotecadigital.puc-campinas.edu.br/tde_busca/arquivo.php?codArquivo=278>. Acesso em: 28 maio 2010.

ESTEBAN NAVARRO, M. A.; GARCÍA MARCO, F. J. Las “primeras jornadas sobre organización del conocimiento: organización del conocimiento e información científica”. **Scire**, Zaragoza, v. 4, p. 149-157, 1995.

ESTIVALET, L. F. **O uso de ícones na Visualização de Informações**. 2000. 90f. Dissertação (mestrado em Ciência da Computação) - Universidade Federal do Rio Grande do Sul, Porto Alegre, 2000.

EVEDOVE, P. R. D. A Política de Tratamento da Informação na Percepção de Catalogadores de Assunto. In: CONGRESSO BRASILEIRO DE BIBLIOTECONOMIA, DOCUMENTAÇÃO E CIÊNCIA DA INFORMAÇÃO, 24., 2011, Maceió. **Anais...** Maceió: CBBBD, 2011. p. 1-12.

FEITOSA, A. **Organização da informação na web**: das tags à web semântica. Brasília, DF: Thesaurus, 2006. 132 p.

FERNEDA, E. **Recuperação da informação**: análise sobre a contribuição da Ciência da Computação para a Ciência da Informação. 2003. 147f. Tese (doutorado em Comunicação) – Escola de Comunicação e Artes, Universidade de São Paulo, São Paulo, 2003.

FRAGOSO, I. da S. **Instituições-memória**: modelos institucionais de proteção ao patrimônio cultural e preservação da memória na cidade de João Pessoa-PB. 2008. 139f. Dissertação (Mestrado em Ciência da Informação)- Universidade Federal da Paraíba, João Pessoa, 2008.

FREITAS, C. M. D.S. et al. Introdução à Visualização de Informações. **RITA**: revista de Informática Teórica e Aplicada, Porto Alegre, v. 8, n. 2, p. 143-158, out. 2001.

FURTADO, C. Educação e bibliotecas digitais. **Revista digital de Biblioteconomia & Ciência da Informação**, Campinas, v. 8, n. 1, p. 103-116, jul./dez. 2010.

FUJITA, M. S. L.; RUBI, M. P.; BOCCATO, V. R. C. O contexto sociocognitivo do catalogador em bibliotecas universitárias: perspectivas para uma política de tratamento da informação documentária. **Datagramazero**: revista de Ciência da Informação, Rio de Janeiro, v. 10, n.2, p. 1-24, abr. 2009.

GALVÃO, M. C. B. Construção de conceitos no campo da Ciência da Informação. **Ciência da Informação**, Brasília, v. 27, n. 1, p. 46-52, jan./abr. 1998.

GARCÍA, A. A. R. Los FRBR y los RDA como campo de oportunidad para alargar la mirada en la organización de la información. In: ENCuentro de Catalogación y Metadatos, 4., 2009, México. **Anais...** México, DF: UNAM/CUIB, 2010. p. 65-84. Disponível em: <http://132.248.242.3/~publica/archivos/libros/iv_encuentro_catalogacion.pdf>. Acesso em: 18 abr. 2013.

GARRIDO ARILLA, M. R. Procedimientos automáticos de creación y transformación de los registros bibliográficos. **Anales de Documentación**, n. 4, p. 127-137, 2001. Disponível em: <<http://revistas.um.es/analesdoc/article/view/2361/2351>>. Acesso em: 18 jan. 2013.

GIL, A. C. **Métodos e técnicas de pesquisa social**. 6. ed. São Paulo: Atlas, 2009.

GONZALES, M.; LIMA, V. L. S. Recuperação de Informação e processamento da Linguagem Natural. In: CONGRESSO DA SOCIEDADE BRASILEIRA DE COMPUTAÇÃO, 23., 2003, Campinas. **Anais...** Campinas: [s.n.], 2003. p.347-395.

GUEDES, R. de M.; SOUZA, R. R. Navegando entre nuvens de etiquetas: uma proposta de Utilização da tag cloud em

catálogos eletrônicos de Bibliotecas. **PontodeAcesso**, Salvador, v. 2, n. 3, p. 2-13, dez. 2008.

GUEDES, V. L. da S. Estudo de um critério para indexação automática derivativa de textos científicos e tecnológicos. **Ciência da Informação**, Brasília, v. 23, n. 3, p. 318-326, set./dez. 1994.

GUIMARÃES, J. A. C. Abordagens teóricas de tratamento temático da informação (TTI): catalogação de assunto, indexação e análise documental. In: GARCÍA MARCO, F. J. Avances y perspectivas en sistemas de información y documentación. Ibersid, Zaragoza: Universidad de Zaragoza, 2009, p. 105-117.

GUIMARÃES, J. A. C. A dimensão teórica do tratamento temático da informação e suas interlocuções com o universo científico da International Society for Knowledge Organization (ISKO). **Revista Ibero-americana de Ciência da Informação**, Brasília, v.1, n.1, p. 77-99, jan./jun. 2008.

GUINCHAT, C.; MENOU, M. Introdução geral às ciências e técnicas da informação e documentação. Brasília: IBICT, 1994.

GRUBER, T. A translation approach to portable ontology specifications. **Knowledge Acquisition**, Amsterdam, v. 5, n. 2, p.199-220, 1993. Disponível em: [≤](#)

<http://tomgruber.org/writing/ontologia-kaj-1993.pdf> > Acesso em: 04 maio. 2013.

INGWERSEN, P. **Information Retrieval Interaction**. London: Taylor Graham, 2002.

INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS. **Guidelines for online public access catalogue (opac) displays**. Munique, 2005.

KOBASHI, N.Y.; SANTOS, R. N. M. dos. Arqueologia do trabalho imaterial: uma aplicação bibliométrica à análise de dissertações e teses. **Encontros Bibli**: revista eletrônica de Biblioteconomia e Ciência da Informação, Florianópolis, n. esp., p. 106-115, 1º sem. 2008.

JOINT STEERING COMMITTEE FOR DEVELOPMENT OF RDA. **RDA: Resource Description and Access**. Canadá, 2006. Disponível em: <<http://www.rda-jsc.org/docs/rda brochure January 2010.pdf>>. Acesso em 07 jul. 2013.

LAGUS, K. et al. WEBSOM for textual data mining. **Artificial Intelligence Review**, v. 13, n. 5-6, p. 345-364, 1999.

LANCASTER, F. W. **Indexação e resumos**: teoria e prática. 2. ed. Brasília: Briquet de Lemos/Livros, 2004.

LARA, M. L. G. de. Linguagem documentária e terminologia. **Transinformação**, Campinas, v. 16, n. 3, p. 231-240, set./dez. 2004.

LE COADIC, Y- F. **A ciência da informação**. 2. ed. Brasília: Briquet de Lemos/Livros, 2004.

LIMA, G. A. B. O. **Mapa Hipertextual (MHTX)- um modelo para organização hipertextual de documentos**. 2004. 199f. Tese (Doutorado em Ciência da Informação) - Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2004. Disponível em: <<http://www.gercinalima.com/glima/uploads/tesegercina.pdf> >. Acesso em: 25 maio 2013.

LUNARDI, M. S. **Visualização em nuvens de texto como apoio à busca exploratória na web**. 2008. 126f. Dissertação (Mestrado em Design) - Escola Superior de Desenho Industrial, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2008.

MACKINLAY, J. D.; ROBERTSON, G. G.; CARD, C. The perspective wall: detail and context smoothly integrated. In: CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS AND GRAPHICS INTERFACE, 9., 1991, New Orleans. **Atas...** New Orleans: ACM-Press, 1991. p. 173–179.

MACULAN, B. C. M. dos S.; LIMA, G. A. B. de O.; PENIDO, P. Taxonomia facetada como interface para facilitar o acesso à

informação em bibliotecas digitais. **Revista ACB**, Florianópolis, v. 16, n. 1, p. 234-249, jan./jun. 2011.

MARTINHO, N. O.; FUJITA, M. S. L. La catalogación de materias: apuntes históricos sobre su normalización. **Scire: Representación y organización del conocimiento**, Zaragoza, v. 16, p. 61-70, 2010.

MAZZA, R. **Introduction to information visualization**. New York, NY: Springer, 2009.

MCGARRY, K. **O contexto dinâmico da informação: uma análise introdutória**. Brasília, DF: Briquet de Lemos/Livros, 1999. 203 p.

MEDEIROS, G. M. **Organização da informação em repositórios digitais: implicações do auto-arquivamento na Representação da informação**. 2010. 273f. Dissertação (Mestrado em Ciência da Informação) Universidade Federal de Santa Catarina, Santa Catarina, 2010.

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

MONTEIRO, F. de S. Organização da informação: proposta de elementos de arquitetura da informação para repositórios. In: ROBREDO, J.; BRÄSCHER, M. (Org.). **Passeios pelo**

bosque da informação: estudos sobre a representação e organização da informação e do conhecimento – eroic. Brasília, DF: IBICT, 2010. p. 130 -145.

MOREIRA, M. P.; MOURA, M. A. Construindo tesouros a partir de tesouros existentes: a experiência do TCI – Tesouro em Ciência da Informação. **DataGramaZero:** revista de Ciência da Informação, Rio de Janeiro, v. 7, n. 4, p. 1-16, ago. 2006. Disponível em: <http://www.dgz.org.br/ago06/Art_01.htm>. Acesso em: 08 maio 2013.

MOREIRA, V. L.; ROMÃO, L. M. S. O funcionamento discursivo das nuvens de tags na rede eletrônica: sentidos sobre Capitolina. **DataGramaZero:** revista em Ciência da Informação, Rio de Janeiro, v.10, n.6, p. 1-17, nov./dez. 2009. Disponível em: < http://www.dgz.org.br/dez09/Art_02.htm >. Acesso em: 15 jun. 2010.

MORENO, F. P. **Requisitos Funcionais para Registros Bibliográficos - FRBR:** um estudo no catálogo da Rede Bibliodata. 2006. 202 f. Dissertação (mestrado em Ciência da Informação) - Departamento de Ciência da Informação e Documentação, Universidade de Brasília, Brasília, 2006.

NASCIMENTO, H. A. D.; FERREIRA, C. B. R. Uma introdução à visualização de informações. **VISUALIDADES**, Goiânia, v.9, n.2, p. 13-43, jul./dez. 2011.

NASCIMENTO, H. A. D.; FERREIRA, C. B. R. Visualização de informações: uma abordagem prática. In: CONGRESSO DA SOCIEDADE BRASILEIRA DE COMPUTAÇÃO, 25., 2005, São Leopoldo. **Anais...** São Leopoldo: UNISINOS, 2005. p. 1262 -1312.

NAVES, M. M. L. Estudo de fatores interferentes no processo de análise de assunto. **Perspectivas em ciência da informação**, Belo Horizonte, v. 6, n. 2, p. 189 - 203, jul./dez. 2001.

NAVES, M. M. L.; KURAMOTO, H. (Org.). **Organização da informação**: princípios e tendências. Brasília, DF: Briquet de Lemos/Livros, 2006. 142p.

NONATO, R. S. **Teoria do conceito e hipertextos**: uma proposta para determinação de relacionamentos em *links* conceituais. 2009. 122 f. Dissertação (Mestrado em Ciência da Informação) – Escola de Ciência da Informação, Universidade Federal de Minas gerais, Belo Horizonte, 2009.

NOVAK, J. D.; CAÑAS, A. J. **The Theory Underlying Concept Maps and How to Construct Them**. Pensacola, FL: Institute for Human and Machine Cognition, 2008. 36 p.
Disponível em: < <http://cmap.ihmc.us/Publications/ResearchPapers/TheoryUnderlyingConceptMaps.pdf>>.
Acesso em: 27 maio. 2013.

OLIVEIRA, E. B.; RODRIGUES, G. M. As concepções de memória na Ciência da Informação no Brasil: estudo preliminar sobre a ocorrência do tema na produção científica. **PontodeAcesso**, Salvador, v. 3, n. 3, p. 216 - 239, dez. 2009.

PINHO, F. A. **Aspectos éticos em representação do conhecimento**: em busca do diálogo entre Antonio García Gutiérrez, Michèle Hudon e Clare Beghtol. 2006. 123 f. Dissertação (Mestrado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2006.

PINHO, F. A. **Aspectos éticos em representação do conhecimento em temáticas relativas à homossexualidade masculina**: uma análise da precisão em linguagens de indexação brasileiras. 2010. 149 f. Tese (doutorado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2010.

PRESSER, N. H.; SILVA, E. L. da; SANTOS, R. N. M. dos. Recursos de Formulação e Visualização de indicadores para apoiar processos de gestão educacional em IESS. **Revista Digital de Biblioteconomia e Ciência da Informação**, Campinas, v. 7, n.2, p. 247-259, jan./jun. 2010.

ROBERTSON, G. G.; CARD, S. K.; MACKINLAY, J. D. Information visualization using 3d interactive animation. **Communications of the ACM**, New York, v. 36, n.4, p. 57-71, 1993.

ROBREDO, J.; CUNHA, M. B. da. **Documentação de hoje e de amanhã**: uma abordagem informatizada da biblioteconomia e dos sistemas de informação. São Paulo: Global, 1994. 400p.

ROWLEY, J. **A biblioteca eletrônica**. Brasília: Briquet de Lemos, 2002.

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

SARACEVIC, T. Interdisciplinary nature of information science. **Ciência da Informação**, Brasília, v.24, n.1, p.36-41, jan./abr. 1995.

SCHOLZ, K. **JavaScript Visual Wordnet**. 2006. Disponível em: <http://kylescholz.com/projects/wordnet/>. Acesso em: 17 jun. 2013.

SILVA, C. G. Considerações sobre o uso de visualização de informação no auxílio a gestão da informação. In: SEMINÁRIO INTEGRADO DE SOFTWARE E HARDWARE, 34., 2007, Rio de Janeiro. **Anais...** Rio de Janeiro: [s.n], 2007. p. 2070-2084.

SILVA, D. L. da. **Uma proposta metodológica para construção de ontologias:** uma perspectiva interdisciplinar entre as ciências da informação e da computação. 2008. 286 f. Dissertação (Mestrado em Ciência da Informação) - Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2008.

SILVA, M. dos R. da; FUJITA, M. S. L. A prática de indexação: análise da evolução de tendências teóricas e metodológicas. **Transinformação**, Campinas, v.16, n. 2, p. 133-161, maio/ago. 2004.

SOUZA, M. I. F.; VENDRUSCULO, L. G.; MELO, G. C. Metadados para a descrição de recursos de informação eletrônica: utilização do padrão Dublin Core. **Ciência da Informação**, Brasília, v. 29, n. 1, p. 93-102, abr. 2000.

SOUZA, R. F. de. A Classificação como Interface da Internet. **DataGramaZero**: revista de Ciência da Informação, Rio de Janeiro, v.2, n.2, p. 1-6, abr. 2000.

SOUZA, R. R. Sistemas de Recuperação de Informação e Mecanismos de Busca na web: panorama atual e tendências. **Perspectivas em Ciência da Informação**, Belo Horizonte, v.11, n. 2, p. 161-173, maio/ago. 2006.

SVENONIUS, E. **The intellectual foundation of information organization**. Cambridge: MIT Press, 2000.

TORRES PONJUAN, D.; PONJUÁN DANTE, G. Influencias ergonómicas en la Visualización de Información. **Informação & Informação**, Londrina, v. 15, n. 1, p. 17-36, jan./jun. 2010.

TRISTÃO, A. M. D.; FACHIN, G. R. B.; ALARCON, O. E. Sistema de classificação facetada e tesauros: instrumentos para organização do conhecimento. **Ciência da Informação**, Brasília, v. 33, n. 2, p. 161-171, maio/ago. 2004.

UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION. **Guidelines for the establishment and development of monolingual thesauri**. Paris, 1971.

VIEIRA, J. M. de L. **Visualização da Informação aplicada à construção de interfaces para recuperação de informação na BDTD-UFPE**. 2011. 61f. Monografia (Bacharelado em Biblioteconomia) – Universidade Federal de Pernambuco, Recife, 2011.

VITAL, L. P.; CAFÉ, L. Proposta para o desenvolvimento de taxonomias em portais corporativos. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 16, n. 4, p. 42-54, 2011.

VITAL, L. P. Taxonomia como elemento estruturante em portais corporativos. **Revista digital de Biblioteconomia e Ciência da Informação**, Campinas, v. 9, n. 2, p. 25-40, 2012.

VITAL, L. P. **Taxonomia como ferramenta para a representação do conhecimento em portais corporativos.** Recife: Ed. Universitária da UFPE, 2010. 87 p.

XAVIER, R. F. **Análise de métodos de produção de interfaces visuais para recuperação da Informação.** 2009. 78 f. Dissertação (mestrado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2009.

XIANG et al. Visualizing criminal relationships: comparison of a hyperbolic tree and a hierarchical list. **Decision Support Systems**, Amsterdam , v. 41, n.2, p. 69–83, nov. 2005.

ZAFALON, Z. R. **Scan for Marc:** princípios sintáticos e semânticos de registros bibliográficos aplicados à conversão de dados analógicos para o formato Marc21 Bibliográfico. 2012. 169 f. Tese (doutorado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2012.

ZHANG, J. **Visualization for information retrieval.** New York: Springer. 2008.

WIVES, L. K. **Estudo sobre agrupamento de documentos textuais em processamento de informação não estruturadas usando técnicas de clustering.** 1999. 102f. Dissertação (Mestrado em Ciência da Computação) - Universidade Federal do Rio Grande do Sul, Porto Alegre, 1999.