

**UNIVERSIDADE FEDERAL DE SANTA CATARINA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA
E GESTÃO DO CONHECIMENTO**

Héctor Andrés Melgar Sasieta

**UM MODELO PARA A VISUALIZAÇÃO DE
CONHECIMENTO BASEADO EM IMAGENS
SEMÂNTICAS**

Florianópolis

2011

Héctor Andrés Melgar Sasieta

**UM MODELO PARA A VISUALIZAÇÃO DE
CONHECIMENTO BASEADO EM IMAGENS
SEMÂNTICAS**

Tese submetida ao Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento para a obtenção do Grau de Doutor em Engenharia e Gestão do Conhecimento.

Orientador: Prof. Dr. Roberto Carlos dos Santos Pacheco

Coorientador: Prof. Dr. José Leomar Todesco e Dr. Fabiano Duarte Beppler

Florianópolis

2011

Catálogo na fonte elaborada pela biblioteca da
Universidade Federal de Santa Catarina

A ficha catalográfica é confeccionada pela Biblioteca Central.

Tamanho: 7cm x 12 cm

Fonte: Times New Roman 9,5

Maiores informações em:

<http://www.bu.ufsc.br/design/Catalogacao.html>

Héctor Andrés Melgar Sasieta

**UM MODELO PARA A VISUALIZAÇÃO DE
CONHECIMENTO BASEADO EM IMAGENS
SEMÂNTICAS**

Esta Tese foi julgada aprovada para a obtenção do Título de “Doutor em Engenharia e Gestão do Conhecimento”, e aprovada em sua forma final pelo Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento.

Florianópolis, 12 de agosto 2011.

Prof. Dr. Paulo Maurício Selig
Coordenador do Curso

Prof. Dr. José Leomar Todesco e Dr. Fabiano Duarte Beppler
Coorientador

Banca Examinadora:

Prof. Dr. Roberto Carlos dos Santos Pacheco
Orientador

Prof. Dr. Tarcisio Vanzin
Membro (UFSC - Universidade Federal de Santa Catarina)

Prof. Dr. Denilson Sell
Membro (UFSC - Universidade Federal de Santa Catarina)

Dr. Frederico Fonseca
Examinador Externo (Penn State University)

Dr. Marcelo Moret
Examinador Externo (CIMATEC)

Dr. Luís Carlos Petry
Examinador Externo (PUCSP - Pontifícia Universidade Católica de
São Paulo)

Dedico este trabalho a minha esposa Mayra e a minhas filhas Daniela e Luciana, minhas fiéis companheiras nesta viagem, minhas razões de viver. Mayra, Daniela, Luciana as amo do fundo de meu coração.

AGRADECIMENTOS

Ao Departamento de Engenharia da Pontifícia Universidad Católica del Perú pela liberação de minhas atividades acadêmicas que permitiu concluir meus estudos no Brasil com sucesso. Em especial aos professores Antonio Montalbetti, Kurt Paulsen y Domingo Gonzalez.

À minha seção de Ingeniería Informática da Pontifícia Universidad Católica del Perú, que me ofereceu todas as facilidades do caso. Em especial aos meus colegas Felipe Solari, Viktor Khlebnikov, Luis Flores, Claudia Zapata, Johan Baldeon e Miguel Guanira.

Ao governo do Brasil por todo o apoio oferecido em especial ao Programa de Estudante-Convênio de Pós-Graduação (PEC-PG), à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), ao Ministério das Relações Exteriores (MRE) e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). Muito obrigado pelo seu apoio financeiro.

Ao Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento (EGC) da Universidade Federal de Santa Catarina (UFSC) em especial ao professor Roberto Carlos dos Santos Pacheco que me recebeu de braços abertos nesta aventura.

Ao Instituto Stela que generosamente prestou suas instalações e a ajuda dos doutores Fabiano Beppler e José Todesco o que teve um efeito catalisador nesta esta pesquisa, fazendo que rapidamente muitos artigos sejam aceitos em eventos nacionais e internacionais.

À família PJ (Procuradoria Jurídica) da Softplan/Poligraph, em especial aos colegas Ésio Cunha, Cristiano Dácio, Juliano Bartelli, Casiano Zago que receberam de mãos abertas a este peruano nos momentos mais críticos na minha estadia no Brasil. A todos os colegas da equipe, obrigado por tudo, a PJ foi a minha segunda casa e fiquei muito feliz de ter sido parte dessa família. Obrigado Rômulo Pelachino, Jorge Turatti, Diego da Silva, Carlos Polegato, Rafael Gomes, William Disner, Gustavo Ângelo, Rafael Dessuy, Fábio Trierveiler e Felipe Durante.

Uma imagem vale mais do que mil palavras

RESUMO

MELGAR SASIETA, Héctor Andrés. **Um modelo para a visualização do conhecimento baseado em imagens semânticas**, 2011. Tese de doutorado em Engenharia e Gestão do Conhecimento - Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento, UFSC, Florianópolis.

Os avanços no processamento e gerenciamento eletrônico de documentos têm gerado um acúmulo grande de conhecimento que tem excedido o que os usuários comuns podem perceber. Uma quantidade considerável de conhecimento encontra-se explicitado em diversos documentos armazenados em repositórios digitais. Em muitos casos, a possibilidade de acessar de forma eficiente e reutilizar este conhecimento é limitada. Como resultado disto, a maioria do conhecimento não é suficientemente explorado nem compartilhado, e conseqüentemente é esquecido em um tempo relativamente curto.

As tecnologias emergentes de visualização e o sistema perceptual humano podem ser explorados para melhorar o acesso a grandes espaços de informação facilitando a detecção de padrões. Por outro lado, o uso de elementos visuais que contenham representações do mundo real que *a priori* são conhecidos pelo grupo-alvo e que fazem parte da sua visão de mundo, permite que o conhecimento apresentado por meio destas representações possa facilmente ser relacionados com o conhecimento prévio dos indivíduos, facilitando assim a aprendizagem.

Apesar das representações visuais terem sido usadas como suporte para a disseminação do conhecimento, não têm sido propostos modelos que integrem os métodos e técnicas da engenharia do conhecimento com o uso das imagens como meio para recuperar e visualizar conhecimento.

Neste trabalho apresenta-se um modelo que visa facilitar a visualização do conhecimento armazenado em repositórios digitais usando imagens semânticas. O usuário, através das imagens semânticas, pode recuperar e visualizar o conhecimento relacionado às entidades representadas nas regiões das imagens. As imagens semânticas são representações visuais do mundo real as quais são conhecidas previamente pelo grupo alvo e possuem mecanismos que permitem identificar os conceitos do

domínio representados em cada região. O modelo proposto apóia-se no *framework* para visualização do conhecimento proposto por Burkhard e descreve as interações dos usuários com as imagens. Um protótipo foi desenvolvido para demonstrar a viabilidade do modelo usando imagens no domínio da anatomia, a *Foundational Model of Anatomy* e a *Unified Medical Language System* como conhecimento do domínio e o banco de dados da *Scientific Electronic Library Online* como repositório de documento.

Palavras-chave: visualização do conhecimento, recuperação do conhecimento, imagens semânticas, mapeamentos semânticos

ABSTRACT

MELGAR SASIETA, Héctor Andrés. **Um modelo para a visualização do conhecimento baseado em imagens semânticas**, 2011. Tese de doutorado em Engenharia e Gestão do Conhecimento - Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento, UFSC, Florianópolis.

Advances in processing and electronic document management have generated a great accumulation of knowledge that is beyond what ordinary users can understand. A considerable amount of knowledge is explained in various documents stored in digital repositories. In many cases, the ability to efficiently access and reuse this knowledge is limited. As a result, most knowledge is not exploited or shared, and therefore it is forgotten in a relatively short time.

The emerging technologies of visualization and the human perceptual system can be exploited to improve access to large information spaces facilitating the patterns detection. Moreover, the use of visual elements that contain representations of the real world that are known a priori by the target group and that are part of his world view, allows that the knowledge presented by these representations can be easily related to their prior knowledge, thereby facilitating learning.

Despite visual representations have been used to support knowledge dissemination, no models have been proposed to integrate knowledge engineering methods and techniques with the use of images as a medium to retrieve and display knowledge.

This work presents a model that aims to facilitate the visualization of the knowledge stored in digital repositories using semantic images. Through the semantic images, the user can retrieve and visualize the knowledge related to the entities represented in the image regions. The semantic images are visual representations of the real world which are known in advance by the target group and have mechanisms to identify domain concepts represented in each region. The proposed model is based on the framework for visualization of knowledge proposed by Burkhard and describes the interactions of users with the images. A prototype was developed to demonstrate the feasibility of the model

using archetypes in the field of anatomy, using the Foundational Model of Anatomy and the Unified Medical Language System as knowledge domain and the database of the Scientific Electronic Library Online as a document repository.

Keywords: knowledge visualization, knowledge retrieval, semantic image, semantic mapping

LISTA DE FIGURAS

Figura 1	Exemplo de pre-atenção.....	35
Figura 2	Pesquisa dedutiva e indutiva.....	38
Figura 3	Pirâmide metodológica.....	41
Figura 4	Suite de modelos CommonKADS.....	44
Figura 5	Modelo da pesquisa.....	45
Figura 6	Pirâmide metodológica aplicada na pesquisa.....	46
Figura 7	Representação das fases da pesquisa.....	48
Figura 8	Modelo de referência para o processo de visualização... ..	52
Figura 9	Quarteto do Anscombe.....	53
Figura 10	Gráficos de dispersão do quarteto do Anscombe.....	53
Figura 11	Processamento pre-atencional da cor.....	54
Figura 12	Princípios da Gestalt.....	55
Figura 13	Metáforas para representar hierarquias.....	57
Figura 14	<i>Framework</i> para a visualização do conhecimento.....	62
Figura 15	Exemplo de esboços.....	63
Figura 16	Exemplo de mapa conceitual.....	64
Figura 17	Exemplo de mapa mental.....	65
Figura 18	Imagens da guerra do Vietnã.....	66
Figura 19	Exemplo de metáforas visuais.....	67
Figura 20	Exemplo de mapas.....	67
Figura 21	Exemplo de visualizações interativas.....	68
Figura 22	Triângulo semiótico de Peirce.....	71
Figura 23	Modelo da organização em CommonKADS.....	101
Figura 24	Diagrama de seqüência do processo “visualização de imagens”.....	103
Figura 25	Diagrama de seqüência do processo “recuperação do conhecimento”.....	104
Figura 26	Estrutura das imagens semânticas.....	109
Figura 27	Estrutura dos repositórios de documentos.....	110
Figura 28	Exemplo de mapeamento semântico de um documento.....	113
Figura 29	Modelo proposto para a visualização do conhecimento.....	114
Figura 30	Arquitetura do protótipo.....	123

Figura 31 As imagens semânticas e sua integração com o conhecimento do domínio.	127
Figura 32 Ontologia de mapeamento.	127
Figura 33 Esquema de anotações das imagens.	128
Figura 34 Anotações dos documentos digitais.	129
Figura 35 Ontologia da Imagem.	134
Figura 36 Organização da base de dados SciELO.	135
Figura 37 Processos implementados no protótipo.	136
Figura 38 Protótipo para visualização do conhecimento.	137
Figura 39 Visualização dos documentos	139
Figura 40 Tela inicial do protótipo.	143
Figura 41 Tela resultante da busca da imagem pelo termo “coração”.	144
Figura 42 Informação da região que representa à entidade “veia cava inferior”	145
Figura 43 Informação do mapeamento da região que representa à entidade “veia cava inferior”	146
Figura 44 Quantidade de documentos que mencionam à entidade “veia cava inferior”	146
Figura 45 Tela resultante da busca da imagem pelo termo “pulmões”.	147
Figura 46 Filtragem de documentos pelo conceito ‘fumaça de cigarro’.	148
Figura 47 Modelo para visualização do conhecimento proposto por Burkhard (2005c)	149
Figura 48 Modelo para visualização RuleViz proposto por Han e Cercone (2000a) e Han e Cercone (2000b)	150
Figura 49 Modelo de visualização colaborativa proposto por Novak e Wurst (2005)	152
Figura 50 Modelo para a exploração visual de vídeos proposto por Luo et al. (2007)	154
Figura 51 Sistema de recomendação do conhecimento proposto por Li e Lu (2010)	155
Figura 52 <i>Framework</i> para a visualização do conhecimento baseada em <i>dashboards</i> proposto por Mohd, Embong e Zain (2010)	157

LISTA DE TABELAS

Tabela 1	Contexto do problema de pesquisa.	27
Tabela 2	Categorias do conhecimento	33
Tabela 3	Características das pesquisas qualitativa e quantitativa	40
Tabela 4	Caracterização da pesquisa.	46
Tabela 5	Perfil estatístico dos quartetos Anscombe.	52
Tabela 6	Diferenças entre revisões tradicionais e revisões sistemáticas	78
Tabela 7	Fases de uma revisão sistemática.....	78
Tabela 8	Adaptação da abordagem PICOC em uma revisão sistemática	80
Tabela 9	Resultados das buscas para a revisão sistemática na área da Visualização da Informação - pergunta P1.	83
Tabela 10	Classificação das buscas para a revisão sistemática na área da Visualização da Informação - pergunta P1.	84
Tabela 11	Resultados das buscas para a revisão sistemática na área da Visualização da Informação - pergunta P3.	86
Tabela 12	Estudos analisados para a revisão sistemática na área da Visualização da Informação - pergunta P3.	97
Tabela 13	Resultados das buscas para a revisão sistemática na área da Visualização do Conhecimento - Pergunta 1.....	98
Tabela 14	Resultados das buscas para a revisão sistemática na área da Visualização do Conhecimento - Pergunta VC2.	98
Tabela 15	Resultados das buscas para a revisão sistemática na área da anotação de imagens.	98
Tabela 16	Planilha OM-1 CommonKADS.	102
Tabela 17	Planilha OM-2 CommonKADS.	103
Tabela 18	Planilha OM-3 CommonKADS para o processo Visualização de Imagens.....	117
Tabela 19	Planilha OM-3 CommonKADS para o processo Recuperação do Conhecimento.	118
Tabela 20	Aplicação do <i>framework</i> para visualização do conhecimento.	119
Tabela 21	Visão esquemática do modelo.	119

SUMÁRIO

1 INTRODUÇÃO	25
1.1 O PROBLEMA DE PESQUISA	25
1.2 OBJETIVOS	29
1.2.1 Objetivo Geral	29
1.2.2 Objetivos Específicos	29
1.3 JUSTIFICATIVA	29
1.4 INEDITISMO DO TRABALHO	30
1.5 PRESSUPOSTOS DO TRABALHO	32
1.5.1 O conhecimento pode ser armazenado além da mente humana	32
1.5.2 Os seres humanos possuem habilidades específicas que permitem processar representações visuais de forma mais rápida que os textos	34
1.5.3 A apresentação da informação em forma visual facilita a construção dos modelos mentais	34
1.6 ESCOPO DO TRABALHO	36
1.7 FASES DA PESQUISA	37
1.7.1 Componentes Metodológicos	38
1.7.1.1 Paradigmas de Pesquisa	38
1.7.1.2 Revisão Sistemática	39
1.7.1.3 Pirâmide Metodológica	40
1.7.1.4 A Metodologia CommonKADS	42
1.7.2 Modelo da Pesquisa	44
1.7.3 Caracterização da pesquisa	46
1.7.4 Etapas do Projeto de Pesquisa	47
1.8 ORGANIZAÇÃO DO DOCUMENTO	50
2 REVISÃO DA LITERATURA	51
2.1 VISUALIZAÇÃO DA INFORMAÇÃO	51
2.1.1 As habilidades dos humanos para processar imagens	52
2.1.2 Metáforas visuais	56
2.2 VISUALIZAÇÃO DO CONHECIMENTO	58
2.2.1 Visualização da informação vs visualização do conhecimento	59
2.2.2 <i>Framework</i> para a visualização do conhecimento ...	61
2.3 TEORIA DA CARGA COGNITIVA	68
2.4 SEMIÓTICA	70
2.5 HIPERMÍDIA E INFOGRAFIA	72

2.6	WEB SEMÂNTICA	73
2.6.1	Ontologias	74
2.6.2	Anotação semântica	74
3	ESTADO DA ARTE	77
3.1	REVISÃO SISTEMÁTICA	77
3.1.1	Etapa 1: Planejamento da revisão	79
3.1.2	Etapa 2: Execução da revisão	81
3.1.3	Etapa 3: Relatórios e divulgação	81
3.2	VISUALIZAÇÃO DA INFORMAÇÃO	82
3.3	VISUALIZAÇÃO DO CONHECIMENTO	90
3.4	WEB SEMÂNTICA E ANOTAÇÃO SEMÂNTICA DE IMAGENS	94
4	MODELO PROPOSTO	99
4.1	INTRODUÇÃO	99
4.2	APLICAÇÃO DA METODOLOGIA COMMONKADS	100
4.3	MODELO NO CONTEXTO DA VISUALIZAÇÃO DO CONHECIMENTO	105
4.4	DESCRIÇÃO DO MODELO	106
4.4.1	COMPOSIÇÃO	108
4.4.1.1	Usuários	108
4.4.1.2	Imagens semânticas	108
4.4.1.3	Repositório de documentos	109
4.4.1.4	Repositório de conhecimento	111
4.4.1.5	Visualização	111
4.4.2	AMBIENTE	111
4.4.3	ESTRUTURAS	112
4.4.3.1	Visão do mundo dos usuários	112
4.4.3.2	Anotações semânticas	112
4.4.4	MECANISMOS	114
4.4.4.1	Visualização das imagens	114
4.4.4.2	Visualização do conhecimento	115
4.4.4.3	Refinamento	115
4.4.4.4	Visualização dos documentos	115
4.5	CONSIDERAÇÕES FINAIS	115
5	IMPLEMENTAÇÃO DO MODELO	121
5.1	INTRODUÇÃO	121
5.2	ARQUITETURA	122
5.3	CAMADA DE CONHECIMENTO	122
5.3.1	FMA Wrapper	123
5.3.2	UMLS Wrapper	125
5.4	CAMADA DE ANOTAÇÃO	125

5.4.1 Anotações Semânticas das Imagens.....	126
5.4.2 Anotações Semânticas dos Documentos	128
5.5 CAMADA DE SUPORTE	132
5.5.1 Gestão das Imagens	132
5.5.2 Repositório dos Documentos Digitais.....	134
5.6 CAMADA DE RECUPERAÇÃO.....	135
5.7 CAMADA DE VISUALIZAÇÃO	136
5.7.1 Busca de imagens	137
5.7.2 Recuperação do conhecimento	137
5.7.3 Visualização de documentos.....	138
5.7.4 Refinamento das buscas	138
5.8 CONSIDERAÇÕES FINAIS	138
6 DEMONSTRAÇÃO DE VIABILIDADE E ANÁLISE	
COMPARATIVA	143
6.1 APRESENTAÇÃO DO PROTÓTIPO.....	143
6.2 COMPARAÇÃO DE MODELO PROPOSTO COM OUTRAS	
ABORDAGENS DE VISUALIZAÇÃO DO CONHECIMENTO	147
6.2.1 Modelo proposto e o modelo de visualização de	
Burkhard (2005c)	147
6.2.2 Modelo proposto e o modelo de visualização Ru-	
leViz de Han e Cercone (2000a) e Han e Cercone	
(2000b)	149
6.2.3 Modelo proposto e o modelo de visualização cola-	
borativa proposto por Novak e Wurst (2005).....	151
6.2.4 Modelo proposto e o modelo para exploração visual	
de vídeos proposto por Luo et al. (2007)	153
6.2.5 Modelo proposto e o sistema de recomendação do	
conhecimento proposto por Li e Lu (2010)	155
6.2.6 Modelo proposto e o <i>framework</i> para a visualização	
do conhecimento baseada em <i>dashboards</i> proposto	
por Mohd, Embong e Zain (2010)	156
6.3 ANÁLISE DO MODELO POR MEIO DE HEURÍSTICAS ..	157
6.3.1 Modelo proposto e as heurísticas perceptivas e cog-	
nitivas de Zuk e Carpendale (2006)	157
6.3.2 Modelo proposto e o mantra da visualização da in-	
formação de Shneiderman (1996).....	160
6.3.3 Modelo proposto e critério de Freitas et al. (2002) .	161
6.3.4 Modelo proposto e as dez heurísticas de usabilidade	
de Nielsen e Mack (1994)	161
7 CONCLUSÕES E TRABALHOS FUTUROS	165

7.1 PUBLICAÇÕES RELACIONADAS AO TRABALHO DA TESE	165
7.2 CONCLUSÕES	165
7.3 TRABALHOS FUTUROS	169
REFERÊNCIAS	173

1 INTRODUÇÃO

1.1 O PROBLEMA DE PESQUISA

A sociedade do conhecimento pode ser definida como uma nova posição estratégica da nossa sociedade onde as perspectivas econômicas e sociais estão concentradas na exploração das tecnologias emergentes e do conhecimento bem definido, onde infra-estruturas de aprendizagem são os principais veículos para a aplicação do conhecimento. O objetivo é uma sociedade com acesso ao conhecimento e aprendizagem para todos (LYTRAS; SICILIA, 2005, p. 4), onde os conhecimentos e os relacionamentos tornam-se as principais fontes de riqueza (STEWART, 2001).

Esta sociedade é construída sobre a sinergia dos indivíduos, grupos, organizações, redes sociais e comunidades que exploram formas eficazes de disseminação do conhecimento e fluxos de aprendizagem (LYTRAS; SICILIA, 2005, p. 4), onde a maior parte do valor adicionado na produção inteligente moderna não reside sobre o valor do material utilizado ou na entrada de capital e trabalho, mas sim no conhecimento incorporado no produto final (EVERS, 2001, p. 11).

Esta nova sociedade é caracterizada, por um lado, por uma explosão sem precedentes de informações digitalizadas (GREENGRASS, 2000) e disponibilizadas para grandes audiências através de entornos *on-line*, e por outro lado, por uma sociedade cada vez mais exigente e ansiosa para aprender e aproveitar os benefícios destes conhecimentos.

O acesso a fontes de informação relevantes permite que as novas descobertas produzidas na comunidade acadêmica sejam divulgadas a mais pessoas (MENDONÇA et al., 2001), mas os avanços no processamento e gerenciamento eletrônico de documentos têm gerado uma acumulação grande de conhecimento que tem excedido o que os usuários comuns podem perceber (HOU; PAI, 2009). Como ilustração, a busca no Scopus de artigos publicados no ano 2009 que correspondem ao termo *“information retrieval”* retornou 6.017 registros¹. Assumindo uma velocidade de leitura de uma página por minuto e 5 páginas por artigo, levaria aproximadamente 501 horas para ler todos os artigos, deixando pouco tempo para ler os 70.569 artigos publicados antes do ano 2009.

Como resultado disto, muito deste conhecimento não é suficien-

¹Busca executada o dia 18 de agosto do ano 2010 através da URL <http://www.scopus.com>

temente explorado nem compartilhado, e conseqüentemente é esquecido em um tempo relativamente curto. Existe pois, uma grande necessidade de gerenciar este “dilúvio” de informações (LÉVY, 2001) e transformá-lo em conhecimento utilizável (KALFOGLOU, 2007, p. 225-226).

Por outro lado, pesquisas têm demonstrado que os usuários têm muitos problemas para traduzir suas necessidades de informação consumindo muito do seu tempo (GILCHRIST, 2003), envolvendo a identificação de literatura pertinente e leitura de um grande número de documentos (SONDHI et al., 2007; THEODOSIOU; ANGELIS; VAKALI, 2008). Isto em parte é gerado pelo crescente número de fontes de informação disponível para os usuários (HAVERKAMP; GAUCH, 1998). Segundo Aksoy (2005) mesmo em aplicações tradicionais de bibliotecas digitais existem numerosas bases de dados eletrônicas que podem ser pesquisadas, o usuário deve selecionar quais as bases que são mais relevantes para a sua pesquisa, postar a consulta e revisar manualmente o conjunto de resultados para satisfazer sua necessidade de informação.

Segundo Borgman (1986) o uso dos sistemas de recuperação de informação tradicionais envolve a articulação da necessidade de informação do usuário em termos precisos e relações que correspondam à estrutura do sistema onde a busca será executada. Em um ambiente automatizado, o usuário deve aplicar dois tipos de conhecimento: o conhecimento dos aspectos mecânicos da busca (*i.e.*, sintaxe e semântica dos termos da pesquisa) e conhecimento dos aspectos conceituais (o “como e o porquê” da busca – quais as formas de restringir e ampliar os resultados da pesquisa) . Uma vez que os aspectos mecânicos da busca são compreendidos, os usuários normalmente podem obter alguns resultados, mas somente quando os aspectos conceituais são entendidos o usuário pode explorar o sistema totalmente.

Geralmente os sistemas de recuperação de informação fornecem uma ampla variedade de filtros e opções que permitem restringir ou ampliar o espaço de busca ou oferecer caminhos alternativos para as pesquisas, mas segundo Jansen, Spink e Saracevic (2000) estas técnicas avançadas são pouco usadas pelos usuários.

Outro dos fatores que dificulta o acesso às fontes de informação é a escrita dos documentos usando diversas linguagens, tanto de sistemas formais (*i.e.*, inglês, português, espanhol, etc.) assim como o vocabulário usado nos textos (BARRETT et al., 2005, p. 39) tornando a montagem das consultas mais complicada ainda. Os sistemas de recuperação de informação tentam “entender” o conteúdo de cada documento mas encontram um grande desafio na execução desta tarefa, pois geralmente o conteúdo não é estruturado; as informações são ge-

ralmente escritas usando linguagem natural sem restrições; e frequentemente cobrem um amplo espectro de matérias sendo difícil tratar as ambigüidades (MITRA; CHAUDHURI, 2000, p. 143).

As tecnologias emergentes de visualização e o sistema perceptual humano podem ser explorados para melhorar o acesso a grandes espaços de informação (ROBERTSON; CARD; MACKINLAY, 1993). Usando representações visuais os seres humanos podem detectar padrões mais rapidamente do que em uma análise direta dos conjuntos de dados (SPOERRI, 1993) e com pouco esforço mental (AHLBERG; SHNEIDERMAN, 1994), reduzindo assim a ansiedade dos usuários na procura da informação (AHLBERG; WISTRAND, 1995), facilitando a aprendizagem (KELLER et al., 2006).

Por outro lado, o uso de imagens que contêm representações do mundo real, que *a priori* são conhecidas pelo grupo-alvo e fazem parte da sua visão do mundo, permite que o conhecimento apresentado através destas imagens possa facilmente ser relacionado com os conhecimentos prévios dos indivíduos, facilitando a aprendizagem (BURKHARD, 2004, 2005c).

É diante deste cenário (ver tabela 1) que surge a seguinte pergunta que motiva o trabalho: **Como facilitar a visualização do conhecimento armazenado em repositórios digitais usando representações visuais em um domínio específico?** O modelo proposto neste trabalho é uma tentativa de resposta a esta pergunta.

Problemas	Dificuldade dos usuários para ler e analisar as informações disponíveis.
Oportunidades	Grupos e comunidades que compartilham uma mesma visão do mundo.
Contexto	Crescimento exponencial da informação em formato digital.
	Organizações no contexto da sociedade do conhecimento.
Soluções	Integrar as informações dos repositórios e usar imagens do mundo real para ajudar no processo de recuperação do conhecimento.

Tabela 1 – Contexto do problema de pesquisa.

A partir da pergunta de pesquisa formulada, elaboraram-se duas questões subjacentes, as quais direcionam este estudo: **(1) Quais os**

tipos de elementos visuais podem ser usados para disseminar conhecimento? A visualização do conhecimento pode ajudar a responder esta questão. Esta área foca no uso de representações visuais visando facilitar a transferência de conhecimento entre pessoas e/ou grupo de pessoas (EPPLER; BURKHARD, 2004; BURKHARD, 2004, 2005a, 2005c; EPPLER; BURKHARD, 2005). Justamente uma das perguntas chaves levantadas por Burkhard ao propor o *framework* para a visualização do conhecimento é “como o conhecimento pode ser visualizado”? Ele propõe sete tipos de elementos visuais que podem assistir de forma efetiva ao processo de disseminação do conhecimento. Estes elementos são os esboços, diagramas, imagens, mapas, objetos, visualizações interativas e *storytelling* (contar histórias) (BURKHARD, 2004; BURKHARD; MEIER, 2005; BURKHARD, 2005a, 2005b, 2005c; EPPLER; BURKHARD, 2007).

O uso de imagens que contêm representações do mundo real, que *a priori* são conhecidas pelo grupo-alvo e fazem parte da sua visão do mundo, permite que o conhecimento divulgado através destas imagens possa facilmente ser relacionado com os conhecimentos prévios dos indivíduos, facilitando a aprendizagem (BURKHARD, 2004, 2005c). O uso de imagens com estas características são interessantes para a disseminação do conhecimento.

(2) Como se podem representar as imagens a fim de facilitar a visualização do conhecimento? Para que as imagens possam servir como estrutura de apoio ao processo de visualização do conhecimento, estas têm que estar representadas usando formalismos que permitam a sua interação com as bases de conhecimento visando que os computadores possam recuperar as entidades que nela estão sendo representadas, possibilitando assim a recuperação do conhecimento vinculado. Neste ponto as ontologias podem servir de apoio para a representação das imagens.

Studer, Benjamins e Fensel (1998) definem uma ontologia como uma “especificação explícita e formal de uma conceitualização compartilhada”. Ontologias são esquemas de metadados, que fornecem um vocabulário controlado dos termos, todos eles explicitamente definidos e com semântica que pode ser processada pelas máquinas. Desta forma as ontologias ajudam as pessoas e as máquinas a se comunicarem com mais eficácia (BECHHOFER, 2002, p. 4).

1.2 OBJETIVOS

1.2.1 Objetivo Geral

Desenvolver um modelo que, usando estruturas semânticas que representam imagens do mundo real, possibilite a visualização do conhecimento armazenado em repositórios de documentos digitais.

1.2.2 Objetivos Específicos

Visando-se atingir o objetivo principal, alguns objetivos específicos são requeridos, como os que se seguem.

- **O1:** Identificar na literatura as condições em que os seres humanos, ao processarem representações visuais, conseguem associar as entidades representadas nestas estruturas com conhecimentos prévios;
- **O2:** Definir uma estrutura que permita descrever semanticamente as imagens com o intuito de explicitar as entidades do mundo real contidas nestas imagens;
- **O3:** Definir no modelo mecanismos que permitam uma integração semântica das imagens e os repositórios de documentos digitais com o intuito de facilitar a recuperação do conhecimento;
- **O4:** Implementar um protótipo que aplica as definições descritas pelo modelo que permita demonstrar assim a viabilidade da proposta.

1.3 JUSTIFICATIVA

Os sistemas de visualização da informação têm sido muito usados para facilitar a compreensão de grandes volumes de informação (GERSHON; EICK; CARD, 1998; CARD; MACKINLAY; SHNEIDERMAN, 1999; YI et al., 2008; FEKETE et al., 2008) baseando-se principalmente nas habilidades que os seres humanos têm para processar representações visuais. Diversos métodos e técnicas têm surgido nos últimos anos visando facilitar a navegação e exploração da informação (HERMAN; MELANÇON; MARSHALL, 2000; KOSARA; HAUSER; GRESH, 2003; CARVALHO; MARCOS, 2009). Entretanto, não existe uma ligação destes métodos para os

processos da gestão do conhecimento (BURKHARD, 2004; XIAO-YUE, 2009). É essa justamente a lacuna que a visualização do conhecimento tenta preencher (CHEN, 2005). Na visualização do conhecimento questiona-se por que o conhecimento tem que ser visualizado e como deve ser feita a visualização (BURKHARD, 2004; BURKHARD; MEIER, 2005; BURKHARD, 2005a, 2005b, 2005c; EPPLER; BURKHARD, 2007).

Os sistemas de visualização tradicionais geralmente usam metáforas visuais para facilitar a associação da informação visual com o conhecimento do mundo real (YI et al., 2008; ZIEMKIEWICZ; KOSARA, 2008), usando por exemplo árvores (REINGOLD; TILFORD, 1981; BUCHHEIM; JÜNGER; LEIPERT, 2002), círculos concêntricos (BOARDMAN, 2000; TEOH; MA, 2002), *treemap* (JOHNSON; SHNEIDERMAN, 1991; SHNEIDERMAN, 1992; BRULS; HUIZING; WIJK, 2000), entre outros. O problema é que mesmo usando metáforas visuais estas podem não ser evidentes para os usuários (FABRIKANT; MONTELLO; MARK, 2010), sobretudo quando a metáfora não faz parte da sua visão do mundo.

É necessário, portanto, a construção de modelos de visualização que usem representações que pertencem à visão do mundo dos usuários, que permitam compreender o conteúdo destas representações e usem este conhecimento para recuperar os documentos armazenados nos repositórios digitais. Uma característica destas representações é que ao serem construídas considerando a visão do mundo dos usuários, a interpretação das mesmas produz uma baixa carga cognitiva apoiando assim a processo de aprendizagem.

1.4 INEDITISMO DO TRABALHO

O modelo proposto neste trabalho apresenta um elo entre os métodos e técnicas da engenharia do conhecimento (*i.e.*, recuperação da informação, extração da informação) e a visualização do conhecimento. Durante o período de revisão da literatura, não foram encontradas trabalhos correlatos com os mesmos objetivos desta tese ou que abordaram o uso de imagens como meio para recuperar e visualizar o conhecimento, o que caracteriza a originalidade da pesquisa. Para verificar o ineditismo do trabalho foram feitas revisões sistemáticas (estas revisões se apresentam com detalhe no capítulo do estado da arte), uma delas com o objetivo de verificar como a área da visualização da informação tem se relacionado com as abordagens extração da informação e recuperação do conhecimento.

Nesta revisão, após analisar os estudos, verificou-se que, no âmbito

da extração da informação e a recuperação da informação, os métodos e técnicas da visualização foram usados para atingir quatro grandes objetivos: a exploração da informação (SONG, 2009), a visualização da informação extraída (SEIFERT I., 2009; CHRISTIAN; PATRICK; CHRISTOPHE, 2006; FAHMI I., 2007; CHEN A.A., 2005), a visualização da informação recuperada (YOSHINAGA; NOBUHARA, 2010; SU; FENG; MA, 2010; PELTONEN et al., 2010; VENNA J., 2010; CASTRO P., 2009; PELTONEN, 2009; ZHANG; DENG, 2008; RYDBERG-COX J.A., 2004; ZHANG J., 2002) e a visualização do histórico das buscas (LEIDE J.E., 2007).

Outra revisão sistemática foi feita visando revisar os modelos propostos para a visualização do conhecimento. Dentre dos artigos estudados na revisão somente o proposto por Melgar, Beppler e Pacheco (2010a) possui objetivos similares a esta pesquisa, mas este estudo corresponde a um dos resultados preliminares deste trabalho.

Um dos estudos revisados corresponde ao o modelo proposto por Burkhard (2005c). Este é um modelo genérico para a visualização do conhecimento baseado principalmente em visualização complementares. Dentro das visualizações complementares podem ser usadas imagens do mundo real, porém este modelo não incorpora semântica nestas representações visuais, sendo isto um aspecto que o diferencia deste modelo.

Outros modelos têm sido propostos para atender domínios específicos tais como visualização das fases do processo de descoberta do conhecimento (HAN; CERCONE, 2000a, 2000b), resultados das buscas de vídeos de notícias (LUO et al., 2007) e indicadores para as instituições de ensino superior (MOHD; EMBONG; ZAIN, 2010). A principal diferença destes trabalhos para o modelo proposto é o domínio de aplicação, enquanto os trabalhos mencionados anteriormente têm sido concebidos para serem usados em domínios específicos, o modelo proposto tem sido concebido para ser aplicado em domínios genéricos desde que este permita a representação visual dos conceitos do domínio.

Outros modelos têm sido propostos para atender domínio genéricos (LIANG; HUANG, 2010; LI; LU, 2010; NOVAK; WURST, 2005), mas usam metáforas visuais para visualizar os resultados das buscas e/ou refinar as consultas. De todos os estudos revisados somente o proposto por Burkhard (2005c) incorpora as imagens como estrutura para visualizar conhecimento, nos demais modelos se usam metáforas visuais ou diagramas.

A contribuição científica inédita principal deste trabalho consiste na obtenção de um modelo de visualização do conhecimento a partir de repositórios de informação baseado em imagens semânticas. As imagens contêm representações visuais que pertencem à visão de

mundo dos usuários e possuem mecanismos que permitem explicitar o conteúdo destas representações. Estes mecanismos são implementados no modelo por meio de anotações semânticas. As anotações semânticas têm sido usadas principalmente como método para representar conteúdo multimídia, usando tanto informação textual (SU et al., 2009; JR; RUBIN, 2009; O'SULLIVAN; WILSON; BERTOLOTTO, 2010) quanto informação semântica (WANG; LIU; CHIA, 2006), em processos tanto manuais quanto automáticos (ZHOU et al., 2007; WANG et al., 2008; JR; RUBIN, 2009; KEPING; XIAOJIE; YIXIN, 2010; LI et al., 2010). Estas representações têm sido usadas basicamente para montar sistemas de recuperação de informação, mas não têm sido usadas como médio para recuperar conhecimento, aproveitando o conhecimento residente nos mapeamentos.

1.5 PRESSUPOSTOS DO TRABALHO

1.5.1 O conhecimento pode ser armazenado além da mente humana

Diversas abordagens têm sido propostas visando definir o conhecimento. Alguns autores da área das tecnologias de informação e na engenharia do conhecimento abordam a definição do conhecimento, distinguindo-a entre dados e informação (SCHREIBER et al., 1999, p. 4-5). Outros autores definem conhecimento como um estado da mente, um objeto, um processo, uma condição para ter acesso à informação ou uma capacidade (ALAVI; LEIDNER, 2001, p. 7-8).

As diferentes abordagens sobre a definição do conhecimento correspondem com a base epistemológica dos pesquisadores, sugerindo assim, uma estratégia diferente para gerenciar o conhecimento e uma diferente perspectiva do papel que cumprem os sistemas que suportam a gestão do conhecimento (ALAVI; LEIDNER, 2001, p. 11). Segundo Venzin, Krogh e Roos (1998) para pesquisar a fundo a definição do conhecimento, quatro dimensões têm que ser consideradas: a questão, a epistemologia, a aparência e a aplicação. A **questão** visa explicar por que é importante pesquisar o conhecimento dentro da organização. A **epistemologia** visa identificar as bases epistemológicas usadas. Isto é importante, pois alguns conceitos podem tomar diferentes formas dependendo da teoria na qual estão baseadas. A **aparência** procura as diferentes formas que o conhecimento pode tomar (conhecimento tácito, explícito, social, codificado, entre outros). Dependendo da base

epistemológica usada, estas formas podem ter significados diferentes. Finalmente a **aplicação** explora a forma em que o conhecimento é aplicado nas organizações.

Segundo Nonaka (1994) existem duas dimensões de conhecimento dentro das organizações: a tácita e a explícita. A dimensão **tácita** é originada pela ação, experiência e o envolvimento em um contexto específico. A dimensão tácita do conhecimento é conformada por elementos cognitivos (como modelos mentais, crenças, paradigmas e pontos de vista) e por elementos técnicos (como *know-how*, habilidades e destrezas aplicadas a um contexto específico). A dimensão **explícita** do conhecimento é articulada, codificada e comunicada em forma simbólica e/ou pela linguagem natural (ALAVI; LEIDNER, 2001, p. 15).

Nonaka (1994) tem identificado também, em outra dimensão, dois tipos de conhecimento: o conhecimento individual e o conhecimento social (chamado também de conhecimento grupal ou coletivo por alguns autores). O conhecimento **individual** é criado pelo indivíduo e existe no indivíduo. Por outro lado, o conhecimento **social** é criado e é inerente às ações coletivas e as interações dos indivíduos atuando como grupo (ALAVI; LEIDNER, 2001, p. 17).

Segundo Lam (2000, p. 492), estas dimensões configuram quatro categorias de conhecimento diferentes (ver tabela 2): i) conhecimento mentalizado (*embrained knowledge*), ii) conhecimento incorporado (*embodied knowledge*), iii) conhecimento codificado (*encoded knowledge*) e iv) conhecimento embutido (*embedded knowledge*). Em outra dimensão, o conhecimento pode ser classificado como: i) **declarativo** (*know-about*), ii) **procedural** (*know-how*), iii) **causal** (*know-why*), iv) **condicional** (*know-when*) e v) **relacional** (*know-with*) (ZACK, 1999).

	conhecimento individual	conhecimento coletivo
conhecimento explícito	conhecimento mentalizado (<i>embrained knowledge</i>)	conhecimento codificado (<i>encoded knowledge</i>)
conhecimento tácito	conhecimento incorporado (<i>embodied knowledge</i>)	conhecimento embutido (<i>embedded knowledge</i>)

Tabela 2 – Categorias do conhecimento. Baseado em Lam (2000, p. 491).

Alguns autores afirmam que o conhecimento é algo inerente aos seres humanos (ZELENÝ, 2005), que é criado de forma invisível no cérebro (DAVENPORT; LONG; BEERS, 1998) e adquirido através da ex-

perícia e interação com o ambiente da pessoa (SUNASSEE; SEWRY, 2003), por outro lado alguns autores afirmam que o conhecimento pode ser explicitado além da mente humana (VARELA; THOMPSON; ROSCH, 1991; KROGH; ROOS, 1995; VERA; CROSSAN, 2003; MILTON; CLARKE; SHADBOLT, 2006). Nesta pesquisa considera que o conhecimento pode ser representado usando diversos formalismos e linguagens; e armazenado em repositórios digitais, mas somente a sua dimensão explícita.

1.5.2 Os seres humanos possuem habilidades específicas que permitem processar representações visuais de forma mais rápida que os textos

Segundo Ware (2000) o processo da visualização pode ser dividido em duas etapas. A primeira delas é chamada de processamento pré-atentivo, nela a informação visual é processada pelo olho e pelo córtex visual primário. A segunda etapa é chamada de processamento atento e compreende a busca por detalhes da cena de visualização, esta etapa é realizada seqüencialmente e mais lentamente pelo sistema visual (RODRIGUES, 2007) .

O processamento pré-atentivo refere-se a identificações visuais de padrões por meio de processos inconscientes de percepção que determinam que determinados objetos sejam trazidos a nossa atenção (RODRIGUES, 2007) e muito rápido (TREISMAN, 1985; WOLFE; TREISMAN; HOROWITZ, 2003). Um exemplo do processamento pré-atencional pode ser observado na figura 1, nela pode-se verificar que a contagem da palavra conhecimento é facilitada quando estas palavras estão enfatizadas.

Na revisão da literatura, ao pesquisar a área de visualização da informação descreve-se com mais detalhe este ponto.

1.5.3 A apresentação da informação em forma visual facilita a construção dos modelos mentais

Segundo a teoria da carga cognitiva, os seres humanos possuem uma memória de trabalho ou memória de curto prazo limitada (MILLER, 1956; SWELLER; MERRIENBOER; PAAS, 1998) a qual é usada para armazenar e manipular informação (KIRSCHNER, 2002). Devido a esta limitação, a informação é transferida para uma memória de maior capacidade (KIRSCHNER, 2002) quando esta não é utilizada (BOWER, 1975).

O Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento (EGC) tem como objetivo a pesquisa de novos modelos, métodos e técnicas de engenharia, de gestão e de mídias do conhecimento, para as organizações e para a sociedade em geral.

O objeto de pesquisa do EGC refere -se aos macro-processos de explicitação , gestão e disseminação do conhecimento . Estes incluem os processos de criação , descoberta, aquisição, formalização/ codificação, armazenamento, uso , compartilhamento, transferência e evolução (e.g., observatório do conhecimento).

a

O Programa de Pós-Graduação em Engenharia e Gestão do **Conhecimento** (EGC) tem como objetivo a pesquisa de novos modelos, métodos e técnicas de engenharia, de gestão e de mídias do **conhecimento**, para as organizações e para a sociedade em geral.

O objeto de pesquisa do EGC refere -se aos macro-processos de explicitação , gestão e disseminação do **conhecimento** . Estes incluem os processos de criação , descoberta, aquisição, formalização/codificação , armazenamento, uso , compartilhamento, transferência e evolução (e.g., observatório do **conhecimento**).

b

Figura 1 – Exemplo de pre-atenção. A contagem das ocorrências da palavra conhecimento na imagem (a) é mais fácil quando os símbolos estão enfatizados, como na imagem (b).

Esta memória é a memória de longo prazo. Quando a informação que se encontra na memória de longo prazo é requerida esta é recuperada para a memória de trabalho. A teoria da carga cognitiva se preocupa pela carga na memória de trabalho durante a execução de uma atividade. Mas a carga da memória de trabalho é afetada tanto pela natureza intrínseca do material e pela maneira em que o material é apresentado (SWELLER; MERRIENBOER; PAAS, 1998).

A teoria da carga cognitiva estabelece orientações para apoiar a apresentação da informação de forma a minimizar a sobrecarga do sistema cognitivo, eliminando assim qualquer carga evitável na memória de trabalho, melhorando assim o desempenho intelectual (MAYER; MORENO, 2002; KIRSCHNER, 2002). Se busca, sobretudo reduzir a carga extrínseca, ou seja, a forma em como são apresentadas as informações, visando que o usuário tenha mais capacidade cognitiva livre para o processo de construção e processamento dos modelos mentais (HÖFFLER; LEUTNER, 2007).

Experimentos têm demonstrado que a compreensão é mais provável de ocorrer quando as necessidades dos usuários têm correspondentes representações visuais e verbais na memória de trabalho ao mesmo tempo (MAYER; MORENO, 2002, 2003; ZHU; WATTS, 2010). Além disso,

a construção de imagens mentais pode reduzir a carga cognitiva e aumentar a capacidade de compreensão (LEUTNER; LEOPOLD; SUM-FLETH, 2009; SCHWAMBORN et al., 2011).

Na revisão da literatura, ao pesquisar a teoria da carga cognitiva descreve-se com mais detalhe este ponto.

1.6 ESCOPO DO TRABALHO

Nesta pesquisa propõe-se um modelo que visa facilitar a visualização do conhecimento a partir de repositórios de documentos digitais baseado em imagens semânticas. Para conceber o modelo foram pesquisadas as áreas de visualização da informação, visualização do conhecimento, ontologia, anotações semânticas, semiótica, infografia e hipermídia.

No modelo proposto tanto as imagens quanto os repositórios de documentos têm sido enriquecidos com conteúdo semântico possibilitando a interoperabilidade entre eles. O conteúdo semântico permite “entender” quais os conceitos representados tanto nas imagens quanto nos documentos facilitando assim a recuperação de conhecimento de forma independente à linguagem da escrita dos documentos e a recuperação dos documentos relacionados aos conceitos representados nas imagens. Desta forma, para recuperar documentos, o usuário só precisa selecionar uma região da imagem sem necessidade de ingressar os termos das buscas como nos sistemas de recuperação tradicionais. Os resultados das buscas são apresentados na própria imagem alterando a coloração das regiões onde se concentra a maior quantidade de documentos ou apresentando, de forma textual, a quantidade de documentos recuperados associados a cada região da imagem.

No decorrer do trabalho discutem-se as habilidades que os humanos têm para processar efetivamente as representações visuais e como estas representações podem ser usadas para transferir conhecimento; contudo, a validação destes pressupostos não é o foco desta pesquisa.

Um dos componentes do modelo é o repositório do conhecimento. Neste repositório são armazenadas as representações do conhecimento usadas para descrever as imagens e os documentos. O objetivo deste componente é armazenar representações do conhecimento existentes, entretanto, não é parte desta pesquisa o desenvolvimento de nenhuma ontologia, taxonomia, meta-tesauro de algum domínio de conhecimento em específico.

Os mapeamentos semânticos são os mecanismos usados no mo-

delo para descrever tanto as imagens quanto os documentos dos repositórios. Estes mapeamentos relacionam as entidades descritas nas imagens e documentos com conceitos representados em algum formalismo e armazenados no repositório do conhecimento. O modelo não define como devem ser implementados os mapeamentos. Ao desenvolver sistemas baseados em conhecimento a partir do modelo proposto, o engenheiro do conhecimento tem a liberdade de escolher a forma de implementar os mapeamentos podendo usar para isto arquivos semi-estruturados, banco de dados, ontologias, etc.

A recuperação do conhecimento é o mecanismo usado pelo modelo para obter o conhecimento armazenado nos documentos dos repositórios. Este mecanismo é usado no processo de visualização alterando a coloração das regiões das imagens a fim de facilmente reconhecer as regiões onde existe maior concentração de documentos associados. Entretanto, esta pesquisa não está focada em métodos nem metodologias para recuperação do conhecimento.

1.7 FASES DA PESQUISA

Esta seção tem como objetivo descrever os procedimentos metodológicos usados na execução deste projeto de pesquisa. Esta seção está estruturada da seguinte maneira: após a descrição dos componentes metodológicos, apresenta-se o modelo da pesquisa, posteriormente se apresenta a caracterização da pesquisa, e finalmente são apresentados as etapas do projeto.

Na seção componentes metodológicos, descreve-se brevemente o método dedutivo e indutivo e sua relação com o paradigma quantitativo e qualitativo. Posteriormente se explica o processo de revisão sistemática que pode ser usado nas pesquisas bibliográficas qualitativas. Descreve-se também a pirâmide metodológica onde a metodologia é vista como um sistema em camadas. Finalmente se revisam os principais elementos da metodologia CommonKADS, uma metodologia da engenharia do conhecimento moderna.

No modelo da pesquisa, se explica como se integram as revisões sistemáticas, a pirâmide metodológica e a metodologia CommonKADS dentro do projeto de pesquisa. Na caracterização da pesquisa, se classifica a pesquisa a partir da sua natureza, a sua abordagem, os seus objetivos e seus procedimentos técnicos. Finalmente, se detalham as etapas do projeto que visam satisfazer os objetivos da pesquisa.

1.7.1 Componentes Metodológicos

1.7.1.1 Paradigmas de Pesquisa

Segundo Gil (1999) o método é “um conjunto de procedimentos intelectuais e técnicos” necessário na pesquisa científica. Como o conhecimento científico fundamenta-se na razão, precisa ser sistemático, a fim de que possa ser testado e comprovado por outros membros da comunidade científica (SANTOS, 2008). Dentre os métodos consolidados da pesquisa científica se encontram o dedutivo e o indutivo (GIL, 1999; LAKATOS; MARCONI, 1993).

O método dedutivo caracteriza-se por chegar a uma conclusão a partir de uma situação geral e genérica para uma particular (SILVA; MENEZES, 2005). Se parte de uma teoria, que define as relações entre conceitos dentro de um conjunto de suposições e restrições fixadas, para formular hipóteses visando à confirmação da teoria. O método indutivo caracteriza-se pelo processo pelo qual, o pesquisador por meio de observações, chega a determinadas conclusões gerais, ou seja, parte-se do específico para o geral (SILVA; MENEZES, 2005). A diferença entre ambos métodos pode ser visualizado na figura 2, nela pode se observar como os processos indutivo e dedutivo são contrários.

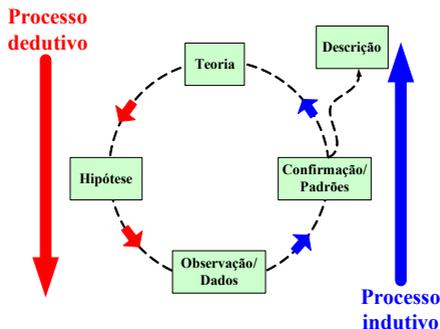


Figura 2 – Pesquisa dedutiva e indutiva.

O método dedutivo e indutivo são usados nos paradigmas quantitativo e qualitativo respectivamente. De um modo geral, existem principalmente dois tipos de modelos ou paradigmas de pesquisa: o qualitativo e o quantitativo (CRESWELL, 2009). Segundo Merriam (1998) na pesquisa qualitativa os pesquisadores estão interessados em compreender o significado que as pessoas construíram, preocupados prin-

principalmente com os processos e não simplesmente com os resultados e os produtos. Procura compreender o fenômeno de interesse a partir da perspectiva do participante e não do pesquisador, sendo o pesquisador o principal instrumento de coleta e análise dos dados. Como foca nos processo, significados e na compreensão, o produto da pesquisa qualitativa é marcadamente descritivo (TREVINOS, 1987). Por outro lado, a pesquisa quantitativa, segundo Creswell (2009), é o processo de investigação de um problema baseado na testagem de uma teoria composta por variáveis, mensuradas com números e analisadas com procedimentos estatísticos. O objetivo na pesquisa quantitativa é desenvolver generalizações que contribuam para o desenvolvimento teórico e que possibilitem uma melhor explicação, compreensão e previsão de um fenômeno (GAUCH, 2003). As diferenças entre ambos os paradigmas podem se apreciar no quadro 3.

1.7.1.2 Revisão Sistemática

Uma revisão sistemática é uma forma de identificar, avaliar e interpretar as pesquisas relevantes disponíveis para um problema de pesquisa específico. As revisões sistemáticas diferem das revisões narrativas tradicionais através da adoção de um processo replicável, científico e transparente (TRANFIELD; DENYER; SMART, 2003, p. 209) que incluem as estratégias para recuperar as evidências a fim que outros profissionais possam reproduzir a revisão (BIOLCHINI et al., 2005, p. 1).

Apesar da infância relativa da revisão sistemática, um razoável consenso emergiu como a suas desejáveis características metodológicas. O *Cochrane Collaboration's Cochrane Reviewers' Handbook* (CLARKE; OXMAN, 2001) e a *National Health Service Dissemination* (KHAN et al., 2001) fornecem uma lista de etapas para a realização de uma revisão sistemática (TRANFIELD; DENYER; SMART, 2003, p. 214).

Entre estas etapas se encontram: i) o **planejamento da revisão** que se divide nas tarefas de identificação das necessidades da revisão, elaboração da proposta para a revisão, e o desenvolvimento de um protocolo de revisão; ii) a **execução da revisão** na qual se realiza a identificação das pesquisas, seleção dos estudos, desenvolvimento de um protocolo de revisão, estudo de avaliação da qualidade, extração de dados e acompanhamento dos progressos, e síntese dos dados; e iii) a **elaboração dos relatórios e divulgação**.

Ponto de comparação	Paradigma Qualitativo	Paradigma Quantitativo
foco da pesquisa	qualitativa	quantitativo
origens filosóficas	fenomenológica, interacionismo simbólico	positivismo, lógica empírica
termos associados	trabalho de campo, etnografia, naturalista, construtivista	experimental, empírica, estatística
objetivos de pesquisa	compreensão, descrição, descoberta, significado, gerador de hipóteses	predição, controle, descrição, confirmação, prova de hipóteses
desenho	flexível, evolvente, emergente	predeterminada, estruturada
amostra	pequena, não randômica, intencional, teórica	grande, randômica, representativa
coleção de dados	pesquisador como principal instrumento, entrevistas, observações, documentos	instrumentos inanimados (escalas, testes, questionários, computadores)
modo de análise	indutivo	dedutivo
resultados	compreensivo, holístico, expansivo, abundantemente descritivo	preciso, numérico

Tabela 3 – Características das pesquisas qualitativa e quantitativa. Adaptado de Merriam (1998).

1.7.1.3 Pirâmide Metodológica

De forma geral uma metodologia pode ser definida como um conjunto de métodos, regras e postulados empregados em uma disciplina (HELJST et al., 1994). Segundo Wielinga, Schreiber e Greef (1989) uma metodologia pode ser considerada como um sistema em camadas, onde cada camada poder vista como um bloco de construção que suporta a camada em cima dele (ver figura 3).

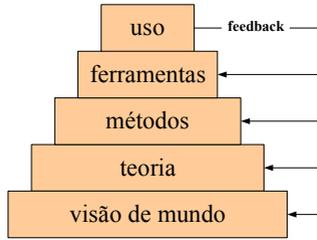


Figura 3 – Pirâmide metodológica.

- A camada de **visão de mundo** da pirâmide refere-se aos princípios e suposições que fundamentam a metodologia estabelecendo assim os limites do escopo. Muitas vezes estes princípios parecem óbvios e até triviais, mas o que pode parecer trivial para uma comunidade poderiam não sê-los para outras. Explicitar a visão do mundo permite estabelecer diálogos com outras metodologias alternativas facilitando sua comparação. Um exemplo de visão de mundo da mecânica newtoniana é a suposição que o mundo é homogêneo (*i.e.*, as mesmas leis podem ser aplicadas em qualquer lugar do universo) (HEIJST et al., 1994). Na metodologias de engenharia do conhecimento modernas (*i.e.*, CommonKADS), a visão de mundo é que a engenharia do conhecimento é uma atividade de modelagem onde padrões permitem o reuso de conhecimento (SCHREIBER et al., 1999). O paradigma orientado a objetos é outro exemplo de visão de mundo no domínio das linguagens de programação (SPEEL et al., 2002).
- A camada da **teoria** sustenta a descrição declarativa do conhecimento disponível no domínio da metodologia. As teorias usualmente são formuladas como um conjunto de teoremas. As leis de Newton (*i.e.*, ação-reação) são parte da capa teórica no domínio da mecânica (HEIJST et al., 1994). No CommonKADS formam parte da teoria as notações textuais e gráficas, as planilhas, as estruturas dos documentos (SCHREIBER et al., 1999).
- A camada dos **métodos** representa às operações que podem ser executadas baseadas na teoria. Em outras palavras a teoria descreve “o quê” e os métodos “o como”. Por exemplo, equações matemáticas, como $x_t = x_0 + v.t$ são usadas como métodos na mecânica newtoniana (HEIJST et al., 1994).
- As **ferramentas** são os dispositivos que suportam os métodos.

Ferramentas no sentido geral da palavra. Termômetros, caneta e papel, martelos e estetoscópios são exemplos de ferramentas (HEIJST et al., 1994).

- O **uso** é a camada de mais alto nível na pirâmide. A metodologia poderá ser usada, por exemplo, em projetos, estudo de casos, etc.

1.7.1.4 A Metodología CommonKADS

Segundo Schreiber et al. (1999) um sistema de gestão do conhecimento deve ser capaz de executar tarefas intensivas em conhecimento ou em todo caso ajudar a executá-las, sendo um fator chave para o sucesso, a capacidade do sistema de lidar com os aspectos organizacionais relevantes. CommonKADS nasce motivada pela necessidade de construir sistemas de qualidade em larga escala de forma estruturada, controlável e repetitiva. Esta metodologia facilita a compreensão do contexto da organização e verifica a necessidade de implantação de um sistema baseado em conhecimento ou alternativas da gestão do conhecimento. A análise da organização é feita através de formulários específicos que tem como objetivo entender e esquematizar a dinâmica das organizações. A experiência acumulada ao longo dos anos tornou o conjunto de modelos CommoKADS em experiência prática dos princípios da análise de conhecimento. Como consequência disso é atualmente a metodologia mais usada em projetos reais (FREITAS, 2003).

Esta metodologia baseia-se em quatro princípios:

1. A engenharia do conhecimento não é uma mineração de conhecimento da cabeça dos especialistas, senão que consistem na construção de modelos sobre os diferentes aspetos do conhecimento humano;
2. A modelagem do conhecimento se concentra primeiramente na estrutura conceitual do conhecimento e deixa os detalhes de programação para depois;
3. El conhecimento possui uma estrutura interna estável que permite analisar e distinguir tipos de conhecimentos específicos;
4. Um projeto de conhecimento deve ser gerenciado tendo em consideração o modelo em espiral. A aprendizagem baseia-se na experiência ao longo do projeto.

CommoKADS integra características metodológicas orientada a modelos abrangendo diversos aspectos de um projeto de desenvolvimento, incluindo análise organizacional, gestão do projeto, aquisição, representação y modelagem do conhecimento; integração e implementação de sistemas. A suíte de modelos CommonKADs tem sua origem e motivação nas seguinte perguntas:

- **Por quê?** Por que um sistema baseado em conhecimento é uma potencial solução? Para que problemas? Que benefícios, custos e impactos se refletiram na organização?
- **Quê?** Qual é a natureza e estrutura do conhecimento envolvido? Qual é a natureza e estrutura da comunicação? A descrição conceptual do conhecimento aplicado nas tarefas é o ponto central aqui;
- **Como?** Como deveria ser implantado o conhecimento em um sistema de computadores? Como se apresentariam as arquiteturas de *software* e mecanismos computacionais? Os aspectos técnicos de infra-estrutura de informação são vistos aqui.

Cada modelo CommonKADS é uma tentativa de resposta a cada uma destas perguntas visando entender o contexto organizacional, os conceptos associados ao conhecimento e os artefatos que possuirão as especificações para construir o sistema baseado em conhecimento. O modelo da organização, o modelo de tarefas e o modelo de agentes analisam o entorno organizacional e os fatores críticos para o sucesso. O modelo de comunicação e o modelo do conhecimento oferecem a descrição conceptual das funções necessárias para resolver o problema e os dados que serão gerenciados. O modelo do projeto transforma todo os demais modelos em uma especificação técnica que serve como base para a implementação do software. Este processo é apresentado na figura 4. É importante notar que não é preciso construir todos os modelos, isto dependerá dos objetivos de cada projeto. Na seqüência, descreve-se brevemente cada um destes modelos.

No modelo da organização se analisam as principais características da organização visando descobrir problemas e oportunidades, estabelecer sua viabilidade e mensurar os impactos que as ações poderiam ocasionar na organização (HOW et al., 1996). No modelo de tarefas se analisa a estruturas das tarefas, suas entradas e suas saídas, pré-condições e critérios de aceitação, assim como as competências e recursos necessários (CHANDRASEKARAN, 1990; SCHREIBER et al., 1993; CHANDRASEKARAN; JOSEPHSON; BENJAMINS, 1998). Os agentes são

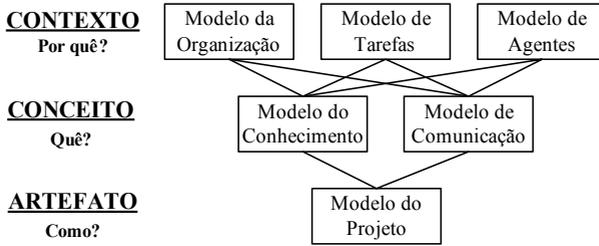


Figura 4 – Suite de modelos CommonKADS.

os executores das tarefas, estes podem ser humanos, sistemas de informação ou qualquer outro tipo de entidade. No modelo de agentes se descreve as principais características dos agentes, em particular sua competência, domínio de execução e restrições. Além disso, é identificada a forma de comunicação entre eles. O propósito do modelo do conhecimento é explicar ao detalhe as estruturas e tipos de conhecimento usado na execução das tarefas. Isto permite uma descrição, independente da plataforma de implantação, do papel que desempenham os diferentes componentes do conhecimento para resolver um problema. Isto faz a este modelo um importante veículo de comunicação entre os usuários e especialistas sobre os aspectos de resolução de problemas. Dado que muitos agentes podem estar envolvidos na execução de uma mesma tarefa, é importante modelar as operações entre eles. Isto é feito no modelo de comunicação, de forma conceitual e independente da sua implementação. Todos os modelos vistos podem ser considerados como constituintes da especificação de requerimentos do sistema baseado em conhecimento. Baseado nestes requerimentos, o modelo do projeto organiza as especificações técnicas do sistema em termos de arquitetura, plataformas de implantação, módulos de *software* e mecanismos computacionais.

1.7.2 Modelo da Pesquisa

O modelo proposto integra o procedimento de revisão sistemática, a pirâmide metodológica e a metodologia CommoKADS. A revisão sistemática foi usada para montar a pirâmide metodológica que suportou as bases para a criação do modelo. O modelo foi criado usando a metodologia CommonKADS. A figura 5 mostra este modelo. Na seqüência, descreve-se como os elementos deste modelo de pesquisa permitem sa-

tisfazer cada um dos objetivos desta pesquisa.

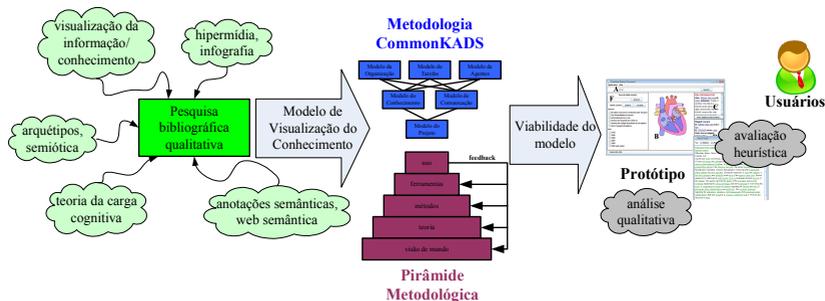


Figura 5 – Modelo da pesquisa.

Visando satisfazer o objetivo 1 (**O1**: *Identificar na literatura as condições em que os seres humanos, ao processarem representações visuais, conseguem associar as entidades representadas nestas estruturas com conhecimentos prévios*), foi feita uma revisão sistemática nos campos da visualização da informação, visualização do conhecimento, arquétipos, semiótica, teoria da carga cognitiva, hipermídia e infografia.

Com o objetivo de satisfazer o objetivo 2 (**O2**: *Definir uma estrutura que permita descrever semanticamente as imagens com o intuito de explicitar as entidades do mundo real contidas nestas imagens*) foi feita também uma revisão sistemática na área de web semântica e anotações de imagens.

Os resultados destas revisões bibliográficas permitiram criar as bases da visão do mundo, teoria e métodos na pirâmide metodológica aplicada na pesquisa. A pirâmide metodológica usada nesta pesquisa é apresentada na figura 6.

A visão do mundo, a teoria e os métodos da pirâmide metodológica da pesquisa foram aplicados à metodologia ComonKADS, visando a execução do objetivo 3 (**O3**: *Definir no modelo mecanismos que permitam uma integração semântica das imagens e os repositórios de documentos digitais com o intuito de facilitar a recuperação do conhecimento*).

Finalmente, com o objetivo de satisfazer o objetivo 4 (**O4**: *Implementar um protótipo que aplica as definições descritas pelo modelo que permita demonstrar assim a viabilidade do modelo proposto*) foi desenvolvido um protótipo, a partir do qual foram feitas análises comparativas qualitativas com outros modelos, avaliações por meio de heurística e avaliações com usuários.

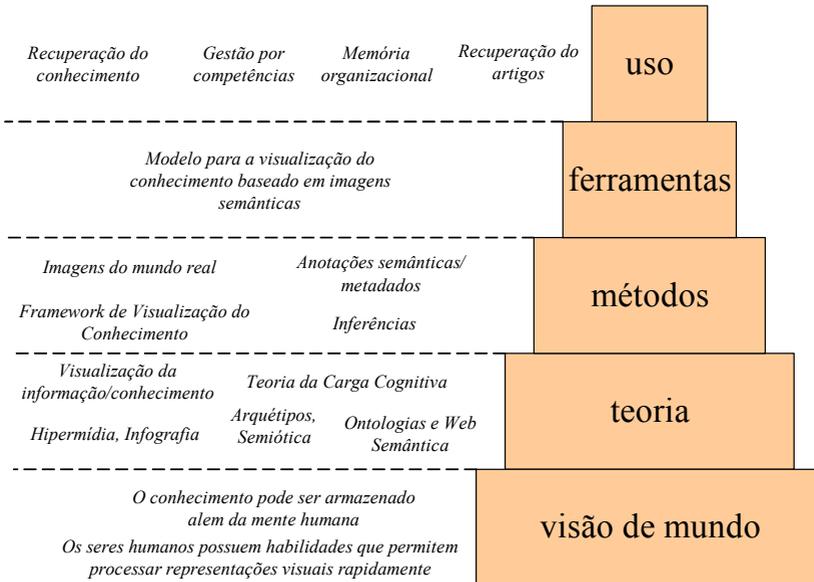


Figura 6 – Pirâmide metodológica aplicada na pesquisa.

1.7.3 Caracterização da pesquisa

Esta pesquisa pode ser classificada como pesquisa aplicada, qualitativa, exploratória e bibliográfica (ver quadro 4):

Ponto de vista	Caracterização
por sua natureza	pesquisa aplicada
por sua abordagem	pesquisa qualitativa
por seus objetivos	pesquisa descritiva
por seus procedimentos técnicos	pesquisa bibliográfica

Tabela 4 – Caracterização da pesquisa.

- A pesquisa pode ser considerada como **aplicada** porque “objetiva gerar conhecimentos para aplicação prática e dirigidos à solução de problemas específicos” (SILVA; MENEZES, 2005). A aplicação prática é a construção de um modelo que permita a visualização de conhecimento usando imagens semânticas que representem a

entidades do mundo real tendo uma baixa carga de trabalho nos processos de busca.

- **Qualitativa** porque foca na compreensão e descrição de como os humanos processam as imagens e como estas podem se relacionar com conhecimento prévio. Também se procura compreender em que condições os humanos executam processo que minimizem sua carga cognitiva.
- **Descritiva** porque “visa descrever as características de determinado fenômeno” (SILVA; MENEZES, 2005). O fenômeno que se procura descrever é como a engenharia do conhecimento suporta a explicitação de imagens com semântica visando à integração destas com repositórios de informação digital a fim de suportar os processos de recuperação de conhecimento usando estas imagens para minimizar a carga de trabalho cognitiva.
- **Bibliográfica** porque é “busca explicar o problema a partir de referenciais teóricos” (CERVO; BERVIAN, 1983). Esta parte da pesquisa se realizará usando revisões sistemáticas com o objetivo de sintetizar os estudos de interesse para esta pesquisa. Os resultados qualitativos obtidos em uma revisão sistemática podem ser apresentados na forma narrativa, quantitativa de estatística de achados qualitativos, ou através da interpretação dos resultados (LOPES; FRACOLLI, 2008).

1.7.4 Etapas do Projeto de Pesquisa

Visando à concepção e ao desenvolvimento da pesquisa, realizou-se uma revisão sistemática sobre como as imagens contendo representações do mundo real podem ser usadas como veículos para visualizar conhecimento. Recorreu-se às áreas de visualização da informação e visualização do conhecimento no intuito de obter um melhor entendimento sobre como os seres humanos processam as imagens e como estas podem ser usadas para divulgar conhecimento. A teoria da carga cognitiva, por seu lado, permitiu estabelecer orientações para apoiar a apresentação da informação de forma a minimizar a sobrecarga do sistema cognitivo, diminuindo a carga na memória de trabalho, melhorando assim o desempenho intelectual. A semiótica forneceu as diretrizes para a organização das imagens semânticas. A semiótica, chamada também de ciência dos signos, pode ser organizada em três categorias: a sintaxe, a semântica e a pragmática. Estas categorias forneceram as

bases para a estruturar as imagens semânticas nos níveis: descritivo, estrutural e semântico. A web semântica forneceu a estruturas de conhecimento necessárias para fazer que o modelo entenda a informação visual contida nas imagens assim como os conceitos mencionados nos documentos textuais facilitando assim a comunicação entre os usuários e o modelo.

Assim, para que os objetivos sejam alcançados, o trabalho foi dividido nas etapas a seguir. Além disso, a figura 7 apresenta, esquematicamente, as fases da pesquisa.

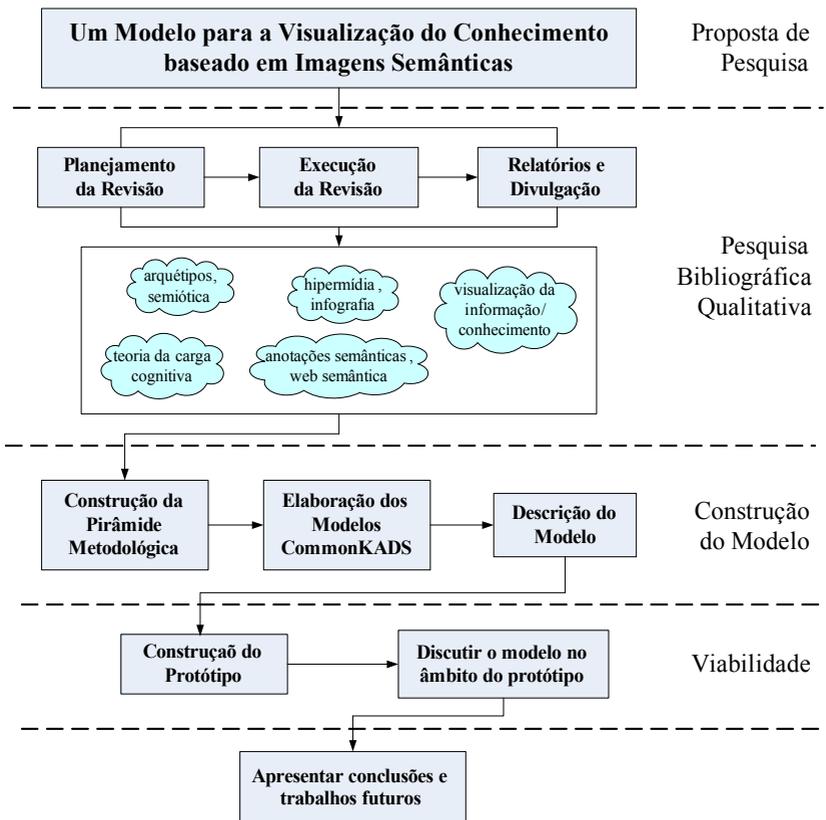


Figura 7 – Representação das fases da pesquisa.

1. **Pesquisa bibliográfica qualitativa:** Nesta etapa, realizou-se um estudo sobre as habilidades inatas que têm os seres humanos

para processar imagens rapidamente a fim de compreender como a visão pode ser usada de forma efetiva para perceber elementos e formas. Também são analisadas as características usadas nos sistemas visualização da informação para explorar os dados e criar novos insights a partir de representações visuais;

Estudaram-se também as diferentes representações visuais que podem ser usadas como veículos para transferir conhecimento. Em particular se analisou o *framework* proposto por Burkhard para a visualização do conhecimento (BURKHARD, 2004, 2005c; BURKHARD; MEIER, 2005; BURKHARD, 2005b, 2005a; EPPLER; BURKHARD, 2007). Este estudo permitiu estabelecer um quadro conceitual que forneceu uma visão geral do campo da visualização do conhecimento e orientou o desenvolvimento do modelo;

Analisaram-se também os métodos existentes na engenharia do conhecimento para descrever tanto as estruturas e entidades representadas nas imagens quanto nos documentos. O foco desta etapa foi encontrar representações que forneçam ao modelo o conhecimento sobre quais as entidades do mundo real estão sendo representadas tanto nas imagens quanto nos documentos. Em particular foi estudado o processo de anotação semântica.

- 2. Construção do modelo:** A partir da revisão bibliográfica se construiu a pirâmide metodológica (ver figura 6) que montou as bases para o desenvolvimento do modelo. A partir desta pirâmide e usando os artefatos da metodologia CommoKADS, trabalhou-se na definição do modelo para visualização do conhecimento. Os estudos sobre visualização da informação e visualização do conhecimento serviram de base para definir as imagens semânticas. Os estudos sobre anotações serviram para criar mecanismos para integrar as imagens e os repositórios de documentos digitais por meio de mapeamentos semânticos. Foram identificados e descritos cada um dos componentes que compõem o modelo.
- 3. Viabilidade:** Nesta etapa foi desenvolvido um protótipo baseado no modelo proposto. O conhecimento a ser visualizado foi recuperado a partir de um repositório de documentos contendo informação referente à área de saúde. As imagens semânticas foram criadas contendo representações de entidades anatômicas. Avaliam-se como os componentes do modelo auxiliam ao usuário no processo de visualização.
- 4. Conclusões e trabalhos futuros:** Apresentam-se as con-

clusões da tese e descrevem-se algumas sugestões de possíveis trabalhos futuros.

1.8 ORGANIZAÇÃO DO DOCUMENTO

Este trabalho é composto, além desta introdução, de mais seis capítulos, como mostrado a seguir:

1. Capítulo 2 - REVISÃO DA LITERATURA: neste capítulo descrevem-se as áreas de pesquisas (*i.e.*, visualização da informação, visualização do conhecimento, anotações) que serviram de embasamento para o trabalho;
2. Capítulo 3 - ESTADO DA ARTE: neste capítulo apresenta-se o estado da arte nos temas de visualização da informação e visualização do conhecimento;
3. Capítulo 4 - MODELO PROPOSTO: neste capítulo apresenta-se o modelo para a visualização do conhecimento;
4. Capítulo 5 - IMPLEMENTAÇÃO DO MODELO: neste capítulo mostra-se a implementação do protótipo a partir do modelo proposto. É apresentada a arquitetura do protótipo, seus módulos e suas relações com o modelo;
5. Capítulo 6 - DEMONSTRAÇÃO DE VIABILIDADE E ANÁLISE COMPARATIVA: neste capítulo discute-se um protótipo baseado no modelo. Também se discute o modelo proposto em comparação com outros modelos propostos para a visualização do conhecimento e se discute o modelo em relação a algumas heurísticas desenvolvidas para a visualização da informação;
6. Capítulo 7 - CONCLUSÕES E TRABALHOS FUTUROS: neste capítulo destacam-se algumas conclusões sobre a pesquisa e descrevem-se sugestões para trabalhos futuros.

2 REVISÃO DA LITERATURA

2.1 VISUALIZAÇÃO DA INFORMAÇÃO

O objetivo principal de um sistema de visualização da informação é representar visualmente informações com o objetivo de facilitar a compreensão dos conjuntos de dados (GERSHON; EICK; CARD, 1998; CARD; MACKINLAY; SHNEIDERMAN, 1999; YI et al., 2008) que, à primeira vista, são volumosos ou não são facilmente de analisar em seu estado original, tornando as informações relevantes em objetos mais naturais, mais semelhantes a outros objetos que já são conhecidos (RAUTER; BENATO, 2005). Os sistemas de visualização da informação são especialmente úteis quando os usuários não sabem que perguntas fazer ajudando-os a restringir rapidamente o espaço da busca e a encontrar dados para estudar mais detalhadamente (FEKETE et al., 2008, p. 2).

Segundo Card, Mackinlay e Shneiderman (1999) a visualização pode ser entendida como o processo de mapeamento de dados a representações visuais para serem apresentadas a um observador humano. Na figura 8 pode-se visualizar um modelo simplificado deste processo onde o fluxo das setas a partir dos dados até o usuário especifica uma série de transformações e o fluxo das setas a partir do humano até as transformações em si, indicam o ajustes por meio de controles operados pelo usuário. As transformações dos dados mapeiam dados crus, não estruturados em dados estruturados; os mapeamentos visuais transformam os dados estruturados em estruturas visuais; e as transformações das visões criam visões a partir das estruturas visuais usando parâmetros gráficos como, por exemplo, a posição e a escala. As interações do usuário controlam os parâmetros destas transformações, por exemplo, restringindo a visualização de certos dados ou mudando a natureza das transformações.

A visualização da informação lida principalmente com dados abstratos, isto é, dados para os quais o usuário não tem nenhum modelo mental preconcebido, por esta razão, a interação é particularmente importante, pois permite explorar e analisar os dados. A interação permite que o usuário implicitamente forme modelos mentais das correlações dos dados, através de reconhecimento de padrões, formando hipóteses mentais e testando-as (KOSARA; HAUSER; GRESH, 2003; CHEN et al., 2009).

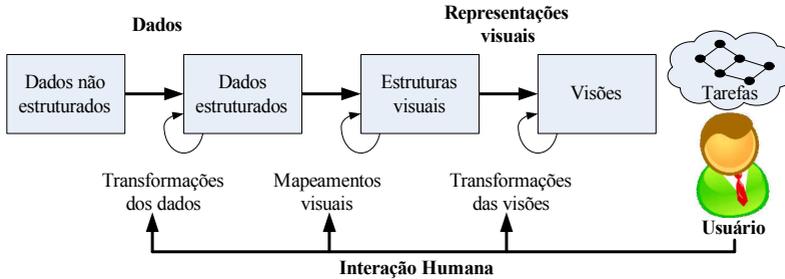


Figura 8 – Modelo de referência para o processo de visualização. Adaptado de Card, Mackinlay e Shneiderman (1999).

2.1.1 As habilidades dos humanos para processar imagens

Os benefícios providos pela visualização parecem depender do fato de ela atuar como um *frame* de referência ou como uma área de armazenamento temporário para os processos da cognição humana. A visualização aumenta a memória dos humanos ao fornecer um extenso conjunto de trabalho para analisar e refletir, e torna-se assim um facilitador externo da cognição (FEKETE et al., 2008). Este benefício pode ser observado ao analisar os gráficos de dispersão do quarteto de Anscombe (ANSCOMBE, 1973). O quarteto de Anscombe compreende quatro conjuntos de dados (ver figura 9) com o mesmo perfil estatístico (ver tabela 5), mas na realidade são muito diferentes quando representados graficamente, como pode ser observado na figura 10. Eles foram criados para demonstrar a importância dos gráficos na análise e os efeitos dos *outliers* nas propriedades estatísticas.

número de observações (n)	11
media de x (\bar{x})	9.0
media de y (\bar{y})	7.5
coeficiente de regressão (b_1)	0.5
equação de regressão linear	$y = 3 + 0.5x$
soma dos quadrados de $(x - \bar{x})$	110.0

Tabela 5 – Perfil estatístico dos quartetos Anscombe. Adaptado de Anscombe (1973).

Segundo Ware (2000) existem principalmente duas teorias na

I		II		III		IV	
x	y	x	y	x	y	x	y
10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89

Figura 9 – Quarteto do Anscombe: quatro conjuntos de dados com valores diferentes e o mesmo perfil estatístico

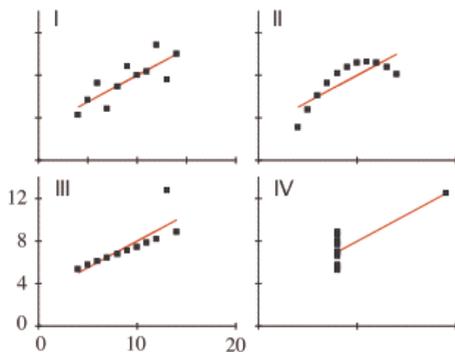


Figura 10 – Gráficos de dispersão do quarteto do Anscombe

psicologia que explicam como a visão pode ser usada efetivamente para perceber elementos e formas. No baixo nível, a teoria do processamento pré-atencional explica que alguns elementos visuais podem ser processados rapidamente e com muita precisão por nosso sistema visual; e no nível mais alto, a teoria da forma ou teoria da Gestalt descreve alguns princípios usados pelo nosso cérebro para entender uma imagem.

O processo de visualização pode ser dividido em duas etapas (TREISMAN, 1985; WOLFE; TREISMAN; HOROWITZ, 2003), na primeira delas, a informação é processada pelo olho e pelo córtex visual primário,

onde neurônios individuais em áreas específicas (denominadas V1, V2, V3, V4, MT) são especializados para identificar características particulares como, por exemplo, orientação, cor, textura, contorno ou movimento. Nesta fase a informação é processada de forma pré-atencional e muito rápido. Posteriormente, na segunda fase, o processamento é dividido dentro de dois subsistemas complementares e independentes, um deles focado na identificação dos objetos (quê) e o outro focado na localização espacial (onde) (BURKHARD, 2005c).

A cor é uma das características que pode ser percebida de forma pré-atentiva. Isto pode ser corroborado ao observar a representação da figura 11-a, onde rapidamente é possível identificarmos o objeto vermelho, neste caso o sistema visual identifica rapidamente este objeto como o elemento fora do padrão da representação. Mas o processamento pré-atentivo só funciona dentro de algumas configurações, na figura 11-b pode-se observar outra representação onde existem uma maior diversidade de cores, apesar de existirem a mesma quantidade de objetos, nenhum padrão pode ser reconhecido pré-atentivamente.

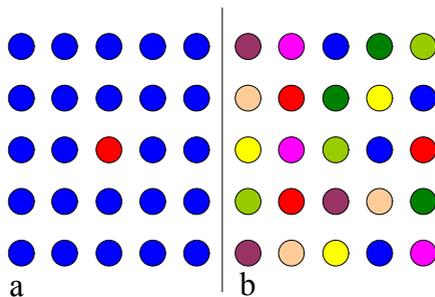


Figura 11 – Processamento pre-atencional da cor. a) representação visual onde a cor pode ser processada de forma pré-atentiva. b) representação visual onde a cor não pode ser processada de forma pré-atentiva.

Os sistemas de visualização da informação são concebidos para fazerem uso do processamento pré-atencional. A cor, por exemplo, tem sido usada como elemento para codificar categorias (FLATLA; GUTWIN, 2010). Existe uma extensa lista de características visuais que podem ser processadas de forma pré-atencional, entre elas podemos mencionar: a orientação da linha, o comprimento ou a largura da linha, o fechamento, a curvatura entre outros.

De acordo com Ware (2000) a teoria da forma ou teoria da Gestalt baseia-se nos seguintes princípios: proximidade, similaridade, con-

tinuidade, simetria, fechamento e tamanho relativo. O princípio da proximidade indica que elementos que estão próximos entre si são percebidos como agrupados. Tal como se pode apreciar na figura 12-a os elementos são percebidos agrupados de acordo com a distância a que se encontram uns dos outros, elementos que estão mais perto de outros tendem a ser percebidos como um grupo pelo que se percebem linhas verticais e horizontais e não círculos. O princípio da similaridade indica que elementos similares tendem a ser agrupados uns com outros como se pode apreciar na figura 12-b onde se percebem linhas verticais e horizontais e não círculos e quadrados. O princípio do fechamento indica que um contorno fechado tende a ser visto como um objeto, este princípio pode ser apreciado na figura 12-c onde se percebe um quadrado e um círculo completos mesmo que estas representações não sejam as desenhadas. O princípio da continuidade indica que elementos visuais que estão levemente conectados ou contínuos tendem a ser agrupados, este princípio pode ser apreciado na figura 12-d onde os pontos próximos à curva se percebem como conectados à mesma. O princípio da simetria indica que dois conjuntos de elementos visuais organizados simetricamente tendem a ser percebidos como um só. O princípio do tamanho relativo indica que pequenos componentes de um padrão tendem a ser percebidos como objetos enquanto os maiores são vistos como segundo plano.

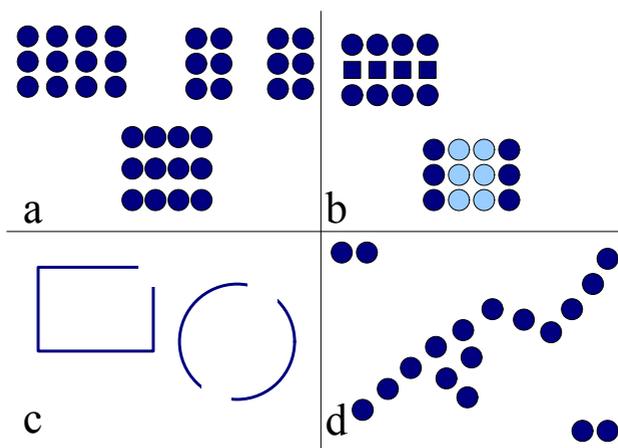


Figura 12 – Princípios da Gestalt: a) proximidade, b) similaridade, c) fechamento d) continuidade.

2.1.2 Metáforas visuais

A visualização da informação contribui de maneira significativa no processo de análise dos dados apoiando a análise exploratória e a análise confirmatória, mas os sistemas de visualização dependem da capacidade dos usuários para interpretar as representações e fazer inferências válidas (MEDDES; MCKENZIE, 2000). De forma geral, os usuários precisam de dois tipos de conhecimento: o conhecimento de como operar o sistema de visualização e o conhecimento do domínio para interpretar as representações visuais (CHEN, 2005). Representações que aproveitam o conhecimento que existe na memória do usuário podem agilizar a construção dos modelos mentais já que quando uma representação visual tem alta consistência com o conhecimento do usuário, evita-se uma carga cognitiva desnecessária para compreender o seu significado (BURKHARD, 2004, 2005c; BAKER; JONES; BURKMAN, 2009), facilitando a compreensão, reconhecimento de padrões e a associação da informação visual com o conhecimento do mundo real. As metáforas visuais facilitam de forma mais efetiva esta associação (YI et al., 2008; ZIEMKIEWICZ; KOSARA, 2008).

Uma metáfora pode ser considerada como o uso de conceitos e conhecimentos a partir de uma área da experiência humana para entender melhor a estrutura, os fenômenos e conceitos de outra que, a princípio, é mais abstrata (AVERBUKH et al., 2007); é a idéia principal que determina o mapeamento do domínio de aplicação para o mundo visual (AVERBUKH, 2001a). Uma das metáforas mais usadas no campo da visualização da informação tem sido a representação de hierarquias. As hierarquias são muitas vezes representadas como árvores.

Na figura 13-a pode-se apreciar a estrutura clássica das árvores. A estrutura da árvore é uma forma de representar a natureza hierárquica de uma estrutura em forma gráfica (REINGOLD; TILFORD, 1981; BUCHHEIM; JÜNGER; LEIPERT, 2002). Neste tipo de estrutura, a árvore pode ser vista como um grafo acíclico dirigido com uma única origem chamada de raiz de tal forma que exista um caminho único dirigido a partir da raiz para qualquer outro nó. A partir desta representação diversas variações são feitas dependendo das necessidades dos sistemas de visualização, por exemplo, a representação da figura 13-b acomoda-se melhor quando há nós com rótulos longos. A visualização radial (figura 13-d) acomoda-se bem em situações onde o foco de atenção encontra-se em um nó colocando este no centro e os outros ao seu redor dispostos em círculos concêntricos (EADES, 1992; BATTISTA et al., 1998; YEE et al., 2001; HEER; CARD; LANDAY, 2005). Os métodos de visualização

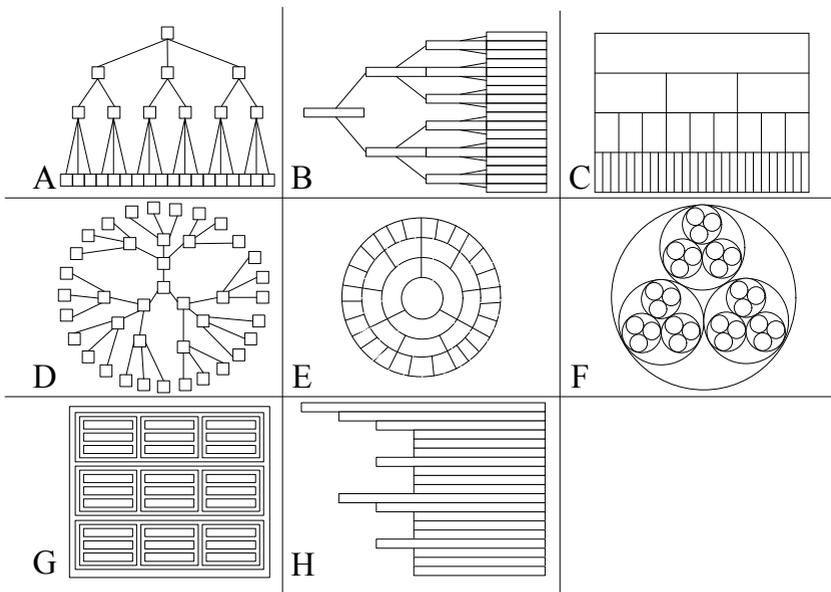


Figura 13 – Metáforas para representar hierarquias. Adaptado de McGuffin e Robert (2010).

com *layout* radial (figura 13-e) acomodam-se bem para revelar grandes estruturas hierárquicas no espaço bidimensional (ANDREWS; HEIDEGGER, 1998; STASKO; ZHANG, 2000; YANG; WARD; RUNDENSTEINER, 2002; DRAPER; LIVNAT; RIESENFELD, 2009). A visualização através de círculos concêntricos (figura 13-f) permite mostrar informações mais contextuais sem comprometer a clareza da área em foco (BOARDMAN, 2000; TEOH; MA, 2002). Os *treemap* (figura 13-g) acomodam-se bem quando o tamanho é a característica mais importante a ser exibida (JOHNSON; SHNEIDERMAN, 1991; SHNEIDERMAN, 1992; BRULS; HUIZING; WIJK, 2000). As listas de árvores (13-h) acomodam-se bem para apresentar informação categorizada.

Representações visuais têm sido usadas amplamente nos sistemas de recuperação e busca de informação por exemplo para melhorar a apresentação dos resultados usando matrizes de ícones (NOWELL et al., 1996), *layouts* dinâmicos (PAEK; DUMAIS; LOGAN, 2004), redes (TENG; CHI, 2007), *treemaps* (EDWARD, 2009), indicadores da relevância dos itens recuperados (VEERASAMY; BELKIN, 1996; HEIMONEN; JHAVERI, 2005; HEIMONEN; SIIRTOLA, 2009), entre outras estruturas. Visualizações também têm sido usadas para ajudar aos usuários na montagem

das consultas por meio de refinamentos iterativos (HOEBER; YANG; YAO, 2007), mapas semânticos, ontologias de domínio, mapas geográficos, *tag clouds* (DORK et al., 2008; HUBMANN-HAIDVOGEL; SCHARL; WEICHSELBRAUN, 2009), *tag clusters* (KNAUTZ; SOUBUSTA; STOCK, 2010); para suportar mecanismos alternativos para explorar documentos (BEDERSON et al., 1996; NGUYEN; ZHANG, 2006); compreender as relações entre os documentos (EBERT et al., 1997; FOWLER; FOWLER, 1998); entre outros.

2.2 VISUALIZAÇÃO DO CONHECIMENTO

A visualização tem sido usada dentro das organizações como meio para explorar informações. Como resultado desta necessidade tem surgido a área da visualização da informação que foca-se na construção de representações visuais de dados abstratos visando facilitar seu entendimento ou ajudar na descoberta de novos *insights*. Nos últimos anos diversos métodos e técnicas da visualização da informação têm sido criados (HERMAN; MELANÇON; MARSHALL, 2000; KOSARA; HAUSER; GRESH, 2003; CARVALHO; MARCOS, 2009), no entanto, não tem sido feita uma ligação destes métodos para o ciclo de vida da gestão do conhecimento (BURKHARD, 2004; XIAO-YUE, 2009). É assim que surge a necessidade da visualização do conhecimento como uma estrutura de mediação para a utilização das diferentes técnicas e métodos de visualização com as tarefas que os gestores de conhecimento têm interesse em executar, começando pela exploração da informação e terminando com a transferência de conhecimento (CHEN, 2005).

A visualização do conhecimento pode ser definida como o uso de representações visuais para melhorar a transferência de conhecimentos entre pelo menos duas pessoas ou grupo de pessoas (EPPLER; BURKHARD, 2004; BURKHARD, 2004, 2005c, 2005a; EPPLER; BURKHARD, 2005). Tornar visível o conhecimento para que ele possa ser acessado, discutido, avaliado, apreciado ou gerenciado, é um objetivo de longa data na gestão do conhecimento, devido a isto a visualização do conhecimento tornou-se recentemente o foco de atenção na comunidade acadêmica e de negócios (XIAO-YUE, 2009). Por outro lado, os avanços no processamento e gerenciamento eletrônico de documentos têm gerado um acúmulo grande de conhecimento nas fontes de informação que tem excedido o que os usuários podem perceber (HOU; PAI, 2009), fazendo que conhecimento relevante permaneça imperceptível dentro de montanhas de informação (SEKI; MOSTAFA, 2007). Diante deste

cenário, a visualização do conhecimento analisa métodos que enfrentam os desafios existentes dentro das organizações referentes à transferência do conhecimento, comunicação para diferentes tipos de destinatário e sobrecarga de informação (BURKHARD; MEIER, 2005).

Atualmente, os conteúdos estão se tornando mais complexos, abstratos e inter-relacionados e principalmente para a transferência de conhecimento, representações verbais têm sido usadas nas organizações. A ignorância sobre as diversas funções das representações visuais leva ao fato de que as visualizações são usadas apenas em domínios específicos como, por exemplo, a publicidade e o marketing. Exemplos de tais funções é a exploração visual interativa dos dados; o mapeamento das informações para apresentar tanto uma visão geral e os detalhes ou o uso de metáforas visuais para transferir e lembrar-se de conceitos complexos. O uso das metáforas visuais, por exemplo, é uma forma eficaz para a transferência de conhecimentos, apresentando diversas vantagens tais como: motivar as pessoas, apresentar novas perspectivas, aumentar a lembrança, apoiar o processo de aprendizagem, chamar a atenção dos usuários e estruturar e coordenar a comunicação (BURKHARD, 2004).

Sob outra perspectiva, a transferência de conhecimento precisa superar outro difícil problema: as diferentes necessidades e perfis dos destinatários. Em geral, os indivíduos só podem entender algo, se isto pode ser conectado a um contexto conhecido. Por isso, conhecer e abordar o conhecimento prévio e o contexto do destinatário individual é determinante. Isso resulta em uma tarefa complexa para a transferência de conhecimento visual: por um lado, os formatos visuais precisam ser específicos a um grupo alvo, o que irá resultar em diferentes formatos e diferentes quantidades de informação descrita; por outro lado, os conteúdos que são apresentados às diferentes partes interessadas não devem ser contraditórios.

2.2.1 Visualização da informação vs visualização do conhecimento

A visualização da informação pode ser definida como o uso de representações visuais e interativas de dados, assistidas pelo computador, visando aumentar a cognição (CARD; MACKINLAY; SHNEIDERMAN, 1999). Seu principal objetivo é representar visualmente informações e facilitar a compreensão de conjuntos de dados que, à primeira vista, são volumosos ou não são facilmente compreendidos em seu estado original (RAUTER; BENATO, 2005). Esta definição encontra-se bem es-

tabelecida, e, como resultado principal cientistas da informação são os envolvidos na comunidade de pesquisa centrada na exploração da informação. O que está faltando, é um discurso sistemático sobre o potencial das visualizações como um meio para a transferência de conhecimentos (BURKHARD, 2004). Esta lacuna é a que a visualização do conhecimento tenta preencher.

Tanto a visualização da informação quanto a visualização do conhecimento exploram as nossas capacidades inatas para processar representações visuais de forma eficaz. Segundo Lengler (2006) as competências visuais podem ser definidas como o conjunto de habilidades que permitem: rapidamente localizar, identificar e avaliar padrões; rapidamente atribuir formas complexas para categorias visuais; estruturar, armazenar e lembrar caminhos de objetos em mapas; re-configurar formas em novos objetos; expressar conceitos com meios visuais em uma ampla variedade de maneiras; construir significados através da integração de diferentes mensagens visuais; imaginar objetos no espaço tridimensional; simular o comportamento futuro dos objetos; e deduzir as regras que regulam os padrões.

Mas a maneira de usar essas competências visuais difere: a visualização da informação visa explorar dados abstratos e criar novos *insights*. A visualização do conhecimento, além disso, visa melhorar a transferência de conhecimentos entre as pessoas. Enquanto a visualização da informação geralmente ajuda a melhorar a recuperação de informação, o acesso e apresentação de grandes conjuntos de dados, e em especial a interação entre humanos e computadores (FEKETE et al., 2008; ZHANG, 2008), a visualização do conhecimento visa essencialmente aumentar comunicação intensiva em conhecimento entre os indivíduos ou grupo de indivíduos (XIAO-YUE, 2009).

A transferência de conhecimento é um processo essencial na gestão do conhecimento e de difícil gerenciamento. Esta acontece em vários níveis: entre indivíduos, de indivíduos para o grupo, entre grupos, entre indivíduos/grupos e a organização; encarando uma série de problemas (*i.e.*, indivíduos com diferentes *backgrounds*, diferentes formações, diferentes motivações, etc.), mas felizmente pode ser abordado usando nossa inata habilidade para processar representações visuais (BURKHARD, 2005c). A visualização do conhecimento oferece soluções para tentar resolver este problema, principalmente fazendo uso de diferentes pressupostos básicos, visíveis e transmissíveis, e provendo *frameworks* visuais que ajudam a estabelecer vínculos entre os diferentes *backgrounds* dos usuários. Desta forma, promove a aprendizagem, a compreensão e a colaboração nas organizações (EPPLER; BURKHARD,

2007). Segundo Zhang (2008) a aplicação da visualização na gestão do conhecimento também pode ser abordada na descoberta do conhecimento; na representações do conhecimento (*i.e.*, regras, classificações, agrupamentos, *outlier*); e na organização da informação (*i.e.*, usando mapas conceituais, mapas cognitivos ou mapas mentais).

2.2.2 *Framework* para a visualização do conhecimento

Nesta seção apresenta-se brevemente o *framework* proposto por Burkhard (BURKHARD, 2004; BURKHARD; MEIER, 2005; BURKHARD, 2005a, 2005b, 2005c; EPPLER; BURKHARD, 2007) para a visualização do conhecimento. Este *framework* baseia-se em cinco perspectivas que respondem a cinco perguntas-chave em relação à visualização de conhecimento:

- Que tipo de conhecimento precisa ser visualizado (conteúdo)?
- Por que o conhecimento deve ser visualizado (propósito, processos da gestão do conhecimento)?
- A quem está sendo destinado (grupo-alvo)?
- Em que contexto deveria ser visualizado (situação: os participantes, local/mídia)?
- Como pode o conhecimento ser visualizado (método, formato)?

A listagem das possíveis respostas a estas perguntas fundamentais leva a um quadro conceitual para representações visuais na gestão do conhecimento que pode fornecer uma visão geral do campo de visualização do conhecimento e orientar a sua aplicação. O quadro apresentado na figura 14, portanto, responde à seguinte pergunta: Por que temos que visualizar o conhecimento em determinada situação e como? Os elementos individuais desta questão são discutidos, na continuação, de forma sucinta.

A **perspectiva do tipo de conhecimento** visa identificar o tipo de conhecimento que precisa ser transferido. O *framework* distingue entre seis tipos de conhecimento: o conhecimento declarativo (*know-what*, por exemplo, fatos) o conhecimento procedural (*know-how*, por exemplo, processos), conhecimento empírico ou experiência (*know-why*, por exemplo, causas), conhecimento relacionado com as pessoas (*know-who*, por exemplo, especialistas), conhecimento baseados na localização (*know-where*, por exemplo, fontes de conhecimento),

TIPO DE CONHECIMENTO	FUNÇÃO	GRUPO ALVO	SITUAÇÃO	TIPO DE VISUALIZAÇÃO
know-what	Criação	A si mesmo	Relatório	Sketches
know-how	Codificação	Outra pessoa	Diálogos	Diagramas
know-why	Transferência	Uma equipe	Discurso	Imagens
know-where	Identificação	Comunidade de prática	Workshop	Mapas
know-who	Aprendizagem	Todos os funcionários	Sistema Especialista	Objetos
know-what-if	Mensuração	Gerentes	Intranet/WWW	Visualizações interativas
Normativa/valores		Público geral	Entorno virtual	Histórias

Figura 14 – *Framework* para a visualização do conhecimento. Adaptado de (BURKHARD, 2005c; EPPLER; BURKHARD, 2007).

conhecimento orientado a contextos (*know-what-if*) e normativa ou conhecimento baseado em valores.

A **perspectiva do propósito da visualização** permite ajudar a distinguir entre as várias razões pelas quais as representações visuais do conhecimento são utilizadas na gestão do conhecimento. Entre os motivos para utilizar a visualização do conhecimento se pode mencionar: o compartilhamento de conhecimento através dos meios visuais, a criação de conhecimento, aprendizagem a partir de representações visuais e a codificação visual das experiências passadas para futuros usuários.

A **perspectiva do grupo-alvo** enfatiza o fato de que a visualização do conhecimento deve acomodar as preferências dos grupos-alvo primários. Limitações do seu conhecimento prévio, suas expectativas e suas restrições de tempo devem ser levadas em consideração ao momento de escolher o formato da visualização. Gerentes, por exemplo, podem sentir-se mais confortáveis analisando um diagrama quantitativo, enquanto que a comunicação com os empregados podem envolver formatos mais simples, como metáforas visuais. No que diz respeito aos efeitos visuais que precisam de atualizações constantes, melhorias ou manutenção (como mapas de conhecimento *on-line*), o grupo responsável por esta tarefa também devem ter em conta, particularmente as suas capacidades tecnológicas. Caso contrário, a visualização do conhecimento pode rapidamente tornar-se obsoleta, devido a manutenção inadequada.

A **perspectiva da situação** salienta o fato de que o uso da visualização depende da configuração física ou virtual e sobre o número de pessoas que interagem para gerenciar o conhecimento. Uma grande comunidade virtual que compartilha sua experiência em portais, precisa

de um tipo diferente de visualização que uma equipe de alta gerência que se reúne a fim de avaliar as distintas capacidades da organização.

A **perspectiva do tipo de visualização** estrutura os métodos de visualização em sete grupos:

Esboços: Os esboços (figura 15) ajudam a visualizar rapidamente uma idéia (EPPLER; BURKHARD, 2004). São usados para assistir tanto a reflexão em grupo quanto ao processo de comunicação fazendo o conhecimento explícito e discutível. Os esboços são fáceis de criar e entender, ajudam a capturar a percepção que o autor quer expressar em um curto espaço de tempo (JIANPING; DA; JIAHUA, 2010).

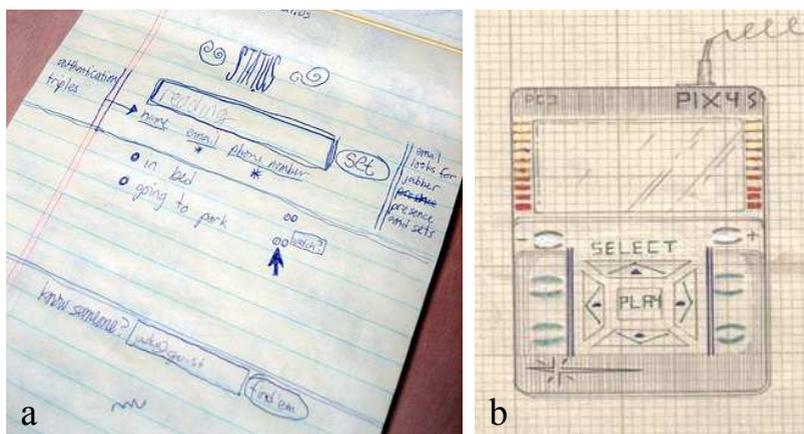


Figura 15 – Exemplo de esboços: a) esboço inicial da idéia do twitter, imagem adaptada de Strange (2007), b) esboço do primeiro aparelho leitor de música digital, imagem capturada da página Web pessoal do Kane Kramer, autor do esboço no ano 1979.

Diagramas: Os diagramas são abstrações, representações esquemáticas usadas para explorar relacionamentos estruturais entre os constituintes. O tipo de conhecimento que é apresentado pelos diagramas é do tipo analítico; diagramas são, portanto, estruturados e sistemáticos. Para a transferência e criação do conhecimento, os diagramas ajudam a fazer conceitos abstratos acessíveis, ajudam a diminuir a complexidade, aumentar a cognição, esclarecer os relacionamentos e a estrutura da informação. Os mapas conceituais e mapas mentais são exemplo deste tipo de representação.

Os mapas conceituais são esquemas gráficos que permitem organizar e representar o conhecimento. O componente fundamental de

um mapa conceitual é a proposição, que consiste em dois conceitos e suas relações (JIANPING; DA; JIAHUA, 2010). Os conceitos nos mapas conceituais são fechados em círculos ou caixas de algum tipo e as relações entre os conceitos são especificadas por uma linha que associa dois conceitos (NOVAK; CAÑAS, 2008). Os mapas conceituais têm sido usados principalmente na educação (ARRUARTE; RUEDA; ELORRIAGA, 2008; CHUNTING; YANG, 2009; DALEY; TORRE, 2010), mas estes podem ser aplicados em outros domínios como ferramenta para geração e comunicação de idéias e estruturas complexas (COFFEY et al., 2002; CAÑAS et al., 2005; TERGAN, 2005; COFFEY; HOFFMAN; CAÑAS, 2006; MORSI; IBRAHIM; WILLIAMS, 2007; CHEN, 2009; XIAO-YUE, 2009; GERKEN; JETTER; REITERER, 2010).

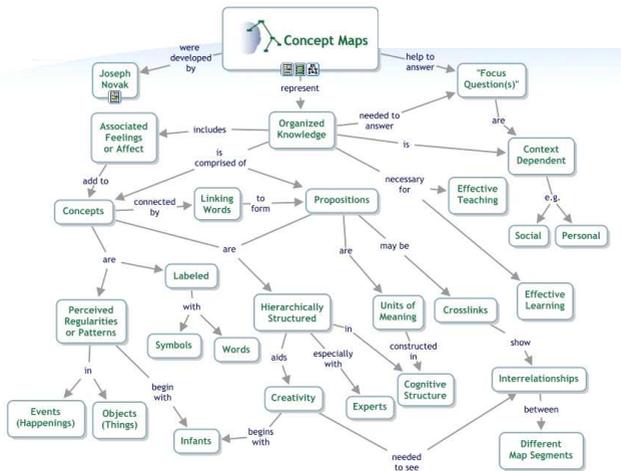


Figura 16 – Exemplo de mapa conceitual. Imagem recuperada do site <http://cmap.ihmc.us/>.

Os mapas mentais (figura 17) são diagramas onde a informação é convertida em uma representação esquemática das palavras-chave mais importantes associadas com um tópico de estudo (FARRAND; HUSSAIN; HENNESSY, 2002), onde graficamente são representadas as conexões entre os conceitos e as idéias que estão relacionadas a um tema central (BUZAN; BUZAN, 1995). Durante a sua produção, uma imagem que representa o tema principal do estudo é inicialmente desenhada no centro e as informações são organizadas através de hierarquias e categorias. Os principais tópicos ou categorias associados com o tema central são captados por ramos que fluem a partir da imagem central

(BUDD, 2004). Os mapas mentais têm sido usados em diversos campos como ferramenta de ensino (BUDD, 2004; WILLIS; MIERTSCHIN, 2006; AYDIN; BALIM, 2009), elemento de apoio em aplicações da ciência da computação (ZUALKERNAN; ABUJAYYAB; GHANAM, 2006; BEEL; GIPP, 2010), projetos de desenho (JIANXIN, 2008) e inclusive para analisar sistemas de governo eletrônico (WU et al., 2009).

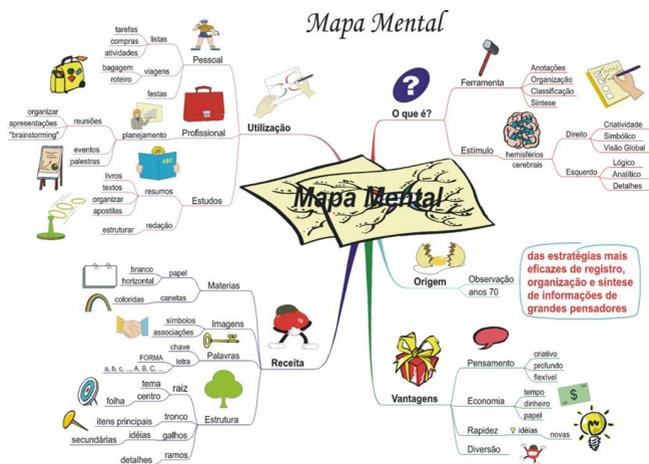


Figura 17 – Exemplo de mapa mental.

Imagens: As imagens abordam as emoções e elas são inspiradoras, atraentes, motivadoras e estimulantes. Assim, elas têm sido usadas como instrumento chave no campo da publicidade. Imagens podem ser aprendidas em menos de um segundo e podem ser recordadas por décadas, como por exemplo, a imagem da figura 18.

O mesmo efeito pode ser usado para transferir conhecimento associado à organização usando, por exemplo, metáforas visuais (AVERBUKH, 2001b). As metáforas (figura 19) podem melhorar a memorização e coordenação de grupos (WORREN; MOORE; ELLIOTT, 2002), mas para isto os usuários devem compreender o que a metáfora representa. Fabrikant, Montello e Mark (2010) sugerem que alguns usuários têm dificuldades em interpretar e compreender metáforas que a princípio são intuitivas para os pesquisadores.

Mapas: Os mapas (figura 20) seguem convenções cartográficas para referenciar o conhecimento. O mapa geralmente consiste de dois elementos uma capa de fundo que representa o contexto e os elementos individuais. No contexto da gestão do conhecimento, os mapas são



Figura 18 – Imagens da guerra do Vietnã. A imagem mostra a Phan Thi Kim Phúc, uma menina vietnamita de nove anos correndo nua devido a que suas roupas foram queimadas por bombas de napalm. A fotografia foi feita pelo Nick Ut em junho do ano 1972, uma imagem que lhe valeu um prêmio Pulitzer. Imagem capturada do site da BBC News através da URL <http://news.bbc.co.uk/2/hi/asia-pacific/4517597.stm>.

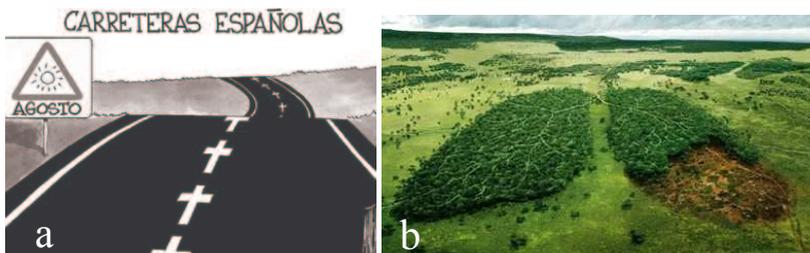


Figura 19 – Exemplo de metáforas visuais: a) *La cruz del tráfico*, metáfora visual que reflete o grande número de acidente de trânsito na Espanha. Imagem capturada do blog Espejo Lúdico recuperada através da URL <http://espejo-ludico.blogspot.com/2007/08/1a-cruz-del-trfico.html>, b) metáfora dos pulmões: as árvores foram posicionadas para parecer pulmões alertando sobre o desmatamento. Imagem capturada do blog SOS rios do Brasil recuperada através da URL <http://sosriosdobrasil.blogspot.com/2009/12/vilmar-berna-portal-mambiente.html>.

chamados de mapas de conhecimento (EPPLER, 2001).

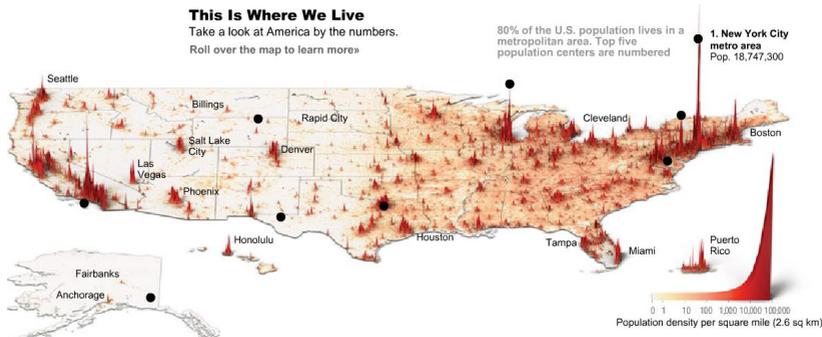


Figura 20 – Exemplo de mapas: a densidade populacional americana sob a forma de picos. Imagem disponível no *Time Magazine* através da URL http://www.time.com/time/covers/20061030/where_we_live/.

Objetos: Objetos no espaço exploram a terceira dimensão. Eles são muito usados para explorar pontos de informação; complementar a visualização tanto por meios físicos quanto por meios digitais; e apresenta os conteúdos desde diferentes pontos de vista.

Visualizações interativas: As visualizações interativas permitem acessar, explorar e dar sentido aos diferentes tipos de informação (LEHMANN; SCHWANECKE; DÖRNER, 2010). Na figura 21 pode-se observar um exemplo de uso de visualizações interativas para sugerir termos de pesquisa.

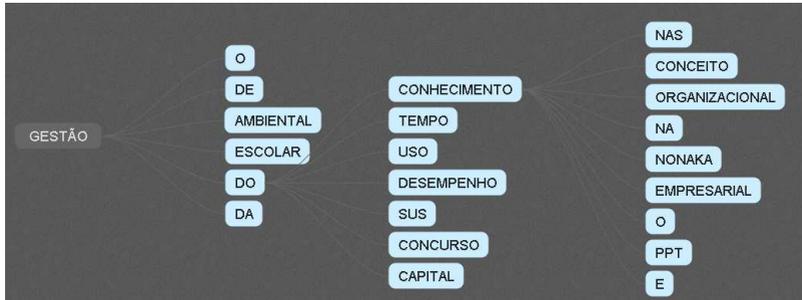


Figura 21 – Exemplo de visualizações interativas. <http://whatdoyousuggest.net/> usa dados do Google para fazer sugestões de termos para pesquisar: ao início o usuário introduz um termo para iniciar a pesquisa, a partir do termo ingressado o sistema apresenta um lista de termos relacionados, ao clicar em qualquer dos termos o sistema apresenta mais termos relacionados.

Histórias: As histórias são representações visuais imaginárias (não são físicas) as quais são eficientes para transferir e disseminar conhecimento através do tempo e do espaço. O uso de histórias é conhecido como *storytelling* e permite transportar uma imagem mental fazendo uso da linguagem falada ou escrita, relevando informações de forma tão eficaz e intuitiva, como se o espectador estivesse assistindo a um filme (GERSHON; PAGE, 2001).

2.3 TEORIA DA CARGA COGNITIVA

Na teoria da carga cognitiva a memória de curto prazo ou memória de trabalho se refere a uma área temporária onde a informação é armazenada e manipulada. É usada para todas as nossas atividades conscientes e é a única memória que podemos controlar (KIRSCHNER, 2002).

De acordo com Miller (1956), a memória de trabalho é muito limitada no que diz respeito ao número de elementos que ela pode

gerenciar, apenas perto de sete itens ou elementos de informação de cada vez. Além disso, como a memória de trabalho é mais comumente usada para processar a informação no sentido de organizar, contrastar, comparar, ou trabalhar sobre a informação de alguma maneira, os seres humanos provavelmente somente são capazes de lidar apenas dois ou três itens simultaneamente (SWELLER; MERRIENBOER; PAAS, 1998).

Devido às limitações da memória de trabalho, a informação é transferida para a memória de longo prazo. A memória de longo prazo é o repositório de conhecimentos e competências permanentes e inclui todas as coisas que não estão sendo utilizadas, mas que são necessárias para a execução de certas tarefas (BOWER, 1975). Alguns cientistas cognitivos acreditam que a capacidade de armazenamento da memória de longo prazo é ilimitada (KIRSCHNER, 2002).

A teoria da carga cognitiva supõe que a memória de trabalho está ligada à memória de longo prazo. A fim de resolver problemas, seções de informação relevante devem ser recuperadas a partir da memória de longo prazo para a memória de trabalho. A facilidade com que as informações possam ser processadas na memória de trabalho é uma das principais preocupações da teoria da carga cognitiva. A carga da memória de trabalho pode ser afetada tanto pela natureza intrínseca do material (carga cognitiva intrínseca) ou, alternativamente, pela maneira em que o material é apresentado (carga cognitiva extrínseca) (SWELLER; MERRIENBOER; PAAS, 1998).

A teoria da carga cognitiva estabelece orientações para apoiar a apresentação da informação de forma a minimizar a sobrecarga do sistema cognitivo, eliminando assim qualquer carga evitável na memória de trabalho, melhorando assim o desempenho intelectual (MAYER; MORENO, 2002; KIRSCHNER, 2002). Ele é baseado em uma arquitetura cognitiva que consiste em uma limitada memória de trabalho, com parcialmente independentes unidades de processamento de informações visuais e de áudio, que interage com uma memória ilimitada de longo prazo (KIRSCHNER, 2002). A carga cognitiva é geralmente considerada uma construção que representa a carga que executa uma determinada tarefa sobre o sistema cognitivo. Ela pode ser conceituada como uma dimensão baseada em tarefas (*i.e.*, a carga mental) e uma dimensão baseada no aprendiz (*i.e.*, o esforço mental), os quais afetam o desempenho (SWELLER; MERRIENBOER; PAAS, 1998).

Se busca reduzir a carga extrínseca (relacionada à forma como são apresentadas as informações), de modo a não carregar ao usuário cognitivamente e ampliar a carga cognitiva relevante para que o usuário tenha mais capacidade cognitiva livre para o processo de construção e

processamento de modelos mentais (HÖFFLER; LEUTNER, 2007). No entanto, reduzir a carga cognitiva não é uma tarefa simples e depende de muitos fatores. Experimentos têm demonstrado que o tipo adequado de estrutura de informação pode depender do nível de especialização (KALYUGA; CHANDLER; SWELLER, 1998; KALYUGA, 2008) e do conhecimento prévio (COOK, 2006).

A teoria da carga cognitiva tem sido usada como suporte nas aplicações da ciência da computação. No campo da multimídia, experimentos baseados na teoria da carga cognitiva sugerem que a compreensão é mais provável de ocorrer quando as necessidades dos usuários têm correspondentes representações visuais e verbais na memória de trabalho, ao mesmo tempo (MAYER; MORENO, 2002, 2003; ZHU; WATTS, 2010). Outros resultados sugerem que a construção de imagens mentais pode reduzir a carga cognitiva e aumentar a compreensão (LEUTNER; LEOPOLD; SUMFLETH, 2009; SCHWAMBORN et al., 2011). Em outro experimento, os resultados indicam que o uso extensivo de *hiperlinks* para comparar e contrastar conceitos quando se leem hipertextos, pode aumentar a carga cognitiva (NIEDERHAUSER et al., 2000).

2.4 SEMIÓTICA

As raízes da semiótica datam de pelo menos da era pré-socrática, onde Hipócrates identificou manifestações do corpo (sinais) como transportadores de mensagens do estado mental e físico. Através de Platão, Aristóteles, Poincaré, Locke, Leibniz e Vico foram discutidos os signos, os símbolos e a comunicação (MICK, 1986). Mas só foi aos inícios do século vinte que a semiótica atingiu sua identificação através dos trabalhos desenvolvidos independentemente pelo filósofo Charles Sanders Peirce e pelo lingüista Ferdinand de Saussure. O termo vem do grego *sêma* (signo). Peirce inicialmente a chamou de *semeioti* e Saussure de *semiology*, mas semiótica é o termo mais comum usado hoje em dia.

De um modo geral, a semiótica analisa as estruturas de produção de significados, tanto verbal como não verbal. Segundo Sebeok (1976) o escopo da semiótica e seu estudo inclui a troca de qualquer mensagem e o sistema de signos que estão subjacentes, sendo o signo o principal conceito.

A semiótica pode tomar duas formas: i) uma semiótica geral que procura responder “Qual é a natureza do significado” e ii) uma semiótica específica que se pergunta “ Como se adquirem significados na nossa realidade (palavras, gestos, mitos, produtos/serviços, teorias,

etc.)?” (RANSDELL, 1977). Para abordar a estas questões, a semiótica investiga o sistema de sinais ou códigos essenciais para todo tipo de comunicação de regras que facilitam a produção de sinais e respostas interpretativas.

Saussure (1916) afirma que desde que a língua era um sistema de sinais, a lingüística deve ser envolvida por uma grande ciência dos signos inserida na sociedade. Saussure definiu ao signo lingüístico como uma relação totalmente arbitrária, entre um conceito que representa o significado (*i.e.*, coração) e um som que representa o significante (*i.e.*, a palavra falada “coração”).

Saussure tentou derrubar a visão tradicional da lingüística onde a língua era vista como uma agregação de distintas unidades (palavras), cada um com um significado diferente. Do ponto de vista de Saussure, as relações e interações entre as palavras têm precedência sobre as palavras individuais quando o significado é formado. Assim, a principal contribuição de Saussure para o estudo dos signos reside na mudança de uma longa ênfase filosófica sobre a natureza das coisas em si mesmas a uma visão de mundo relacional onde o significado deriva das prioridades com que os seres humanos constroem e percebe entre os signos em um sistema.

Para Peirce, a semiótica é o processo de comunicação por qualquer tipo de signo, um signo de qualquer coisa que represente algo (seu objeto), para alguém (o seu intérprete), em algum aspecto (seu contexto). Assim como Saussure, Peirce também explica os processos dos signos em termos de relações, mas Peirce baseia-se em relações triádicas, onde a ação ou influência é envolvida por pela cooperação de três elementos: o signo, o objeto e o interpretante (ver figura 22).

Triângulo semiótico de Peirce

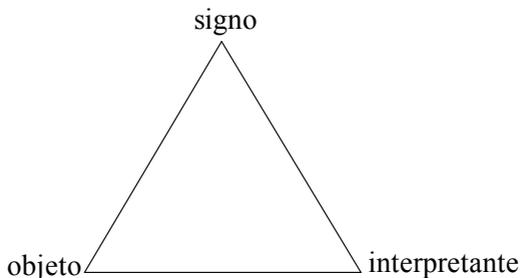


Figura 22 – Triângulo semiótico de Peirce.

Por outro lado, Morris (1946) define a semiótica como a ciência dos signos, onde suas principais divisões são a sintaxe, a semântica e a pragmática. Cada uma destas divisões e a semiótica como um todo podem ser puras, descritivas ou aplicadas. A semiótica pura elabora a linguagem para falar sobre os signos, a semiótica descritiva estuda os signos atuais e a semiótica aplicada utiliza o conhecimento sobre os signos para a realização de diversos propósitos. Assim, a sintaxe é o estudo das relações signo-signo, a semântica estuda as relações signo-objeto e a pragmática estuda as relações signo-interpretante.

2.5 HIPERMÍDIA E INFOGRAFIA

A hipermídia pode ser entendida com uma rede de nós (*i.e.*, artigos, documentos, arquivos, cartões, páginas, quadros, telas) que contém informação (*i.e.*, texto, imagem, vídeo, animação, som) conectada por meio de vínculos ou referências. Onde cada nó pode oferecer diferentes forma de representação (*i.e.*, verbais, pictóricas) através de diferentes modalidades sensoriais (*i.e.*, visual, auditiva) que podem ser usadas de forma interativa. A *World Wide Web* é um exemplo clássico de hipermídia.

O termo hipertexto é mais comumente usado nas aplicações do texto enquanto hipermídia é usado para transmitir a inclusão de outras mídias, especialmente de som e vídeo (SHNEIDERMAN, 1986). A intriga do hipertexto é que ela estende o texto linear tradicional, com a oportunidade de “pular” para vários documentos relacionados, voltar para documentos já revisados, clicar nos índices, nas tabelas de conteúdo, procurar por termos de busca, e usar outras ferramentas de navegação que permitam melhorar profundamente a experiência do leitor (SHNEIDERMAN, 1986).

O desenho de aplicação de hipermídia difere ao desenho de outros *softwares* na medida em que envolve a navegação, bem como da interface do usuário e os problemas de processamento da informação. Um projeto hipermídia envolve capturar e organizar a estrutura de um domínio complexo e torná-lo claro e acessível aos usuários (GARZOTTO; PAOLINI; SCHWABE, 1993).

A infografia por outro lado, pode ser entendida como o uso de representações visuais para a disseminação das informações, dados ou do conhecimento. Estas representação apresentam informações complexas de forma rápida e clara (NEWSOM; HAYNES, 2007). A infografia tradicionalmente é caracterizada por um conjunto de elementos visuais,

como gráficos, mapas ou diagramas que ajuda na compreensão de um determinado conteúdo baseado em texto (RAJAMANICKAM, 2002).

A representação (literalmente, para apresentar de novo) fornece a base para toda a comunicação. Nós podemos transmitir idéias sobre coisas que não estão materialmente em nossa presença apenas por invocar uma representação mental adequada. Essas representações mentais são armazenadas em categorias e são caracterizadas por um grau de abstração. Por exemplo, quando olhamos para uma cadeira, um armário e uma mesa, podemos processá-los como uma ampla classe de objetos (“móveis”), e não como casos individuais da classe. Assim, pensamos em termos de categorias, e as relações entre as categorias. Essa capacidade de generalizar, de formar associações, para detectar relacionamentos, e buscar conforto na organização é a pedra angular do nosso sistema de comunicação. Quando as palavras e frases fazem parte do sistema de comunicação verbal, imagens e representações gráficas formam o sistema de comunicação visual (RAJAMANICKAM, 2002).

2.6 WEB SEMÂNTICA

Bemers-Lee, Hendler e Lassila (2001) definem a Web Semântica como a evolução da Web atual. Ela tem como objetivo enriquecer as páginas Web com conteúdo semântico, criando um ambiente onde agentes de *software* e usuários podem trabalhar de forma cooperativa. A engenharia de ontologias é a tecnologia chave para a Web Semântica, ela integra a compreensão humana simbólica com a capacidade de processamento das máquinas (JOHN; DIETER; VAN, 2003), além disso fornece estruturas para o desenvolvimento de bases de conhecimento (O’LEARY, 1999).

A Web Semântica nasce motivada pelos problemas encontrados na Web atual, destacando-se entre estes a baixa precisão das buscas, o que resulta em um alto número de documentos não relacionados ou na dificuldade para recuperar documentos relevantes. Isto acontece, pois não existe suporte nos mecanismos das buscas mais populares para a análises do significado dos documentos e das palavras claves colocadas nas consultas (ANTONIOU; HARMELEN, 2004). Entre outras necessidades não relacionadas às buscas na Web podemos mencionar: a grande quantidade de dados nos sistemas de informação, a dificuldade de integrar conteúdo de diferentes repositórios de informação e a falta de integração semântica entre os componentes de *software* (JOHN; DIETER; VAN, 2003; STUDER et al., 2000; DACONTA; SMITH; OBRST, 2003).

2.6.1 Ontologias

As ontologias são esquemas de metadados que fornecem um vocabulário controlado de termos, todos eles explicitamente definidos e com semântica que pode ser entendida pelos computadores. Desta forma as ontologias ajudam às pessoas e as máquinas a se comunicar com mais eficiência (BECHHOFFER, 2002).

Uma ontologia pode tomar uma variedade de formas, mas necessariamente incluirá um vocabulário de termos e alguma especificação do seu significado. Isto inclui definições, uma indicação explícita da forma em que os termos estão inter-relacionados coletivamente no domínio e as possíveis restrições na interpretação dos termos (USCHOLD et al., 1998). Studer, Benjamins e Fensel (1998) definem ontologia como uma “especificação explícita e formal de uma conceitualização compartilhada”.

O conhecimento nas ontologias pode ser especificado através de cinco componentes: conceitos, relações, funções, axiomas e instâncias (CORCHO; PEREZ, 2000; GRUBER, 2003). O conceito é representado por meio de classes e suas respectivas propriedades (STAAB et al., 2001; SURE; STUDER, 2002), as propriedades de cada conceito descrevem seus atributos e suas características. Entre tanto existem alguns conceitos que compreendem subclasses e outros que são definidos por mais de uma classe. Neste caso, a relação entre as classes estabelece um tipo de interação entre conceitos que é fundamental para o seu entendimento (CHANDRASEKARAN; JOHNSON; BENJAMINS, 1999). Uma função é um tipo especial de relação que consideram um argumento. Os axiomas são utilizados para criar restrições sobre as informações descritas nos conceitos e verificar a validade do conhecimento. As instâncias, também chamadas de indivíduos, é o termo usado para materializar um conceito (BEPPLER, 2008).

2.6.2 Anotação semântica

A anotação pode ser definida como o processo de tornar explícita a interpretação de um documento ou uma parte dele (MEENA; KUMAR; ROMARY, 2004). As pessoas têm usado as anotações de diversas maneiras como, por exemplo, para lembrar-se de uma parte interessante ou importante em um documento (MARSHALL, 1997, 1998; WOLFE, 2000); como suporte para resumir o conteúdo (O'HARA; SELLEN, 1997); para fazer comentários e sugestões (KAHAN et al., 2002; MARSHALL; BRUSH, 2004); e inclusive para apoiar o trabalho colaborativo (LELAND;

FISH; KRAUT, 1988; SPRING et al., 1997; CADIZ; GUPTA; GRUDIN, 2000; WENG; GENNARI, 2004), e a discussão colaborativa (CONKLIN; BEGEMAN, 1988; MASHAYEKHI; FEULNER; RIEDL, 1994; DAVIS; HUTTENLOCHER, 1995; GINSBURG, 1998; VATHANOPHAS; SPRING, 2001; PERRY; AGARWAL; MCPARLAND, 2002). Anotações de documentos no formato digital, chamadas de anotações digitais, são semelhantes às anotações em papel, mas proporcionam benefícios adicionais que as anotações em papel não podem oferecer: as anotações digitais não estão limitadas fisicamente; podem ser feitas de forma simultânea; e podem ser armazenadas separadas do documento, sem precisar alterar a estrutura ou o conteúdo do documento (SANNOMIYA et al., 2001; RAMACHANDRAN; KASHI, 2003).

Segundo Marshall (1998) existem dois tipos de anotações: as anotações informais e as anotações formais. As anotações informais são o tipo de anotações que as pessoas fazem nos documentos impressos. As anotações formais por outro lado, são metadados que seguem certos padrões que obedecem a algumas especificações que definem a sua estrutura, a sua semântica, a sua sintaxe, e mesmo os valores que as anotações podem assumir. Os metadados são dados sobre os dados, é uma parte secundária da informação que visa descrever a informação principal (SRIVASTAVA; VELEGRAKIS, 2007). Os metadados podem ser definidos como um esquema formal de descrição de recursos, aplicando-se a qualquer tipo de objeto, seja este digital ou não (ORGANIZATION, 2004). Ao ser um esquema formal, os metadatos podem ser entendidos e processados pelas máquinas (AGOSTI; FERRO, 2007; HASLHOFER; KLAS, 2010).

A criação de metadados é uma das principais técnicas usadas para anotar documentos (STEFFEN; ER; SIEGFRIED, 2001; KIRYAKOV et al., 2004). Os metadados podem ser atribuídos a uma ampla variedade de documentos, pode ser expressa em diversas linguagens e vocabulários (CORCHO, 2006), e pode ser feita de forma manual, automática ou semi-automática (UREN et al., 2006). Ontologias têm sido usadas para anotar documentos (STEFFEN; ER; SIEGFRIED, 2001; CORCHO, 2006). Estruturas ontológicas acrescentam valor às anotações semânticas permitindo a realização de inferências e a navegação conceptual (STEFFEN; ER; SIEGFRIED, 2001). Quando as ontologias são usadas, os documentos são etiquetados com descrições semânticas (*i.e.*, informações sobre classes e instâncias) (KIRYAKOV et al., 2004).

3 ESTADO DA ARTE

3.1 REVISÃO SISTEMÁTICA

Uma revisão sistemática é uma forma de identificar, avaliar e interpretar as pesquisas relevantes disponíveis para um problema de pesquisa específico, área temática, ou fenômeno de interesse. Os estudos individuais que contribuem para uma revisão sistemática são chamados de estudos primários, uma revisão sistemática é uma forma de um estudo secundário (KITCHENHAM, 2004, p. 1). As revisões sistemáticas diferem das revisões narrativas tradicionais através da adoção de um processo replicável, científico e transparente, em outras palavras, um método detalhado que visa minimizar o viés através de uma pesquisa bibliográfica exaustiva de estudos publicados e não publicados, e fornece uma trilha da auditoria das decisões dos autores, os procedimentos e as conclusões (TRANFIELD; DENYER; SMART, 2003, p. 209), incluindo as estratégias para recuperar as evidências e definindo de forma explícita o foco da pesquisa a fim que outros profissionais possam reproduzir a revisão (BIOLCHINI et al., 2005, p. 1). As principais diferenças entre as revisões tradicionais e as revisões sistemáticas têm sido sintetizadas no quadro 6.

O movimento baseado em evidências teve um grande impacto em determinadas disciplinas. Principalmente foram aplicados na ciência médica, onde o ritmo de produção de conhecimento tornou difícil a sínteses da informação (TRANFIELD; DENYER; SMART, 2003, p. 208). Na última década, a ciência médica tem feito progressos significativos na tentativa de melhorar a qualidade do processo de revisão visando sintetizar as pesquisas de uma forma sistemática, transparente e reproduzível com o objetivo de oferecer subsídios para as tomadas de decisões nas organizações orientadas à saúde (TRANFIELD; DENYER; SMART, 2003, p. 209). Posteriormente outras áreas como a educação, o serviço social e a justiça criminal se aderiram a esta nova abordagem (TRANFIELD; DENYER; SMART, 2003, p. 209).

Apesar da infância relativa da revisão sistemática, um razoável consenso emergiu como a suas desejáveis características metodológicas. O *Cochrane Collaboration's Cochrane Reviewers' Handbook* (CLARKE; OXMAN, 2001) e a *National Health Service Dissemination* (KHAN et al., 2001) fornecem uma lista de etapas (ver tabela 7) para a realização de uma revisão sistemática (TRANFIELD; DENYER; SMART, 2003, p. 214). Estas etapas têm sido adaptadas também para outras áreas como a

Aspecto	Revisões tradicionais	Revisões sistemáticas
Problema	Frequentemente amplo em escopo	Frequentemente focado em perguntas de pesquisa
Identificação da pesquisa	Normalmente não é especificada, potencial viés	Fontes abrangentes e uma explícita estratégia da busca
Seleção	Normalmente não é especificada, potencial viés	Seleção baseado em critérios aplicados uniformemente
Avaliação	Variável	Avaliação crítica e rigorosa
Síntese	Muitas vezes uma síntese qualitativa	Qualitativos e/ou síntese quantitativa
Inferências	Às vezes baseadas em evidências	Usualmente baseadas em evidências

Tabela 6 – Diferenças entre revisões tradicionais e revisões sistemáticas. Adaptado de Mulrow (1998).

engenharia de *software* (KITCHENHAM, 2004, p. 3).

Etapa	Fase
Planejamento da revisão	Identificação das necessidades da revisão Elaboração da proposta para a revisão Desenvolvimento de um protocolo de revisão
Execução da revisão	Identificação das pesquisas Seleção dos estudos Desenvolvimento de um protocolo de revisão Estudo de avaliação da qualidade Extração de dados e acompanhamento dos progressos Síntese dos dados
Relatórios e divulgação	Elaboração do relatório e as recomendações Obter provas na prática

Tabela 7 – Fases de uma revisão sistemática.

3.1.1 Etapa 1: Planejamento da revisão

Os estágios iniciais das revisões sistemáticas, devem ser um processo iterativo de definição, clarificação e refinamento (TRANFIELD; DENYER; SMART, 2003; KITCHENHAM, 2004). É necessário realizar estudos prévios para avaliar a relevância e o tamanho da literatura e para delimitar a área do assunto ou o tema (TRANFIELD; DENYER; SMART, 2003, p. 214). Dentro da ciência médica o pesquisador deve chegar também a uma pergunta de revisão. Na revisão sistemática estas informações são explicitadas através de um documento formal chamado protocolo de revisão. O protocolo de revisão especifica os métodos que serão utilizados em uma revisão sistemática específica (KITCHENHAM, 2004, p. 4), contém informações sobre as questões específicas abordadas pelo estudo, a população (ou amostra), que é o foco do estudo, a estratégia de busca para identificação dos estudos relevantes, e os critérios de inclusão e exclusão dos estudos incluídos na revisão. O resultado da busca de informações deve ser uma lista completa de artigos e documentos em que a revisão será baseada. Somente estudos que atendem todos os critérios de inclusão do protocolo são incorporados na revisão (TRANFIELD; DENYER; SMART, 2003, p. 215).

A especificação das perguntas da pesquisa é a parte mais importante de qualquer revisão sistemática (KITCHENHAM, 2004, 2007). No campo médico se recomendam construir a pergunta da pesquisa baseada em três enfoques: a população, as intervenções, e os resultados. Posteriormente Petticrew e Roberts (2005) sugerem o uso da abordagem PICOC (das siglas em inglês *Population, Intervention, Comparison, Outcome, Context*). Nesta abordagem as pergunta da pesquisa é montada baseada na população, intervenção, comparação, resultado e contexto.

- Uma questão pode se referir a grupos de **população** muito específica, por exemplo usuarios de sistemas, testadores novatos ou experientes engenheiros do conhecimento. Na medicina as populações são definidos de modo a reduzir o número de estudos prospectivos primários;
- A **intervenção** é a metodologia/ferramenta/tecnologia/procedimento que aborda um tema específico, por exemplo, as tecnologias para realizar tarefas específicas, tais como especificação de requisitos, modelagem do conhecimento, ou estimativa de custo de *software*;

- A **comparação** é a metodologia/ferramenta/tecnologia/processo com o qual a intervenção está sendo comparada;
- Os **resultados** devem estar relacionados a fatores de importância para os profissionais, tais como confiabilidade ou custos de produção reduzidos. Todos os resultados relevantes devem ser especificados. Por exemplo, em alguns casos, necessitam de intervenções que melhorem a algum aspecto da produção sem afetar a outra, por exemplo confiabilidade, sem aumento no custo;
- O **contexto** é o lugar onde a comparação ocorre (por exemplo, academia ou indústria), os candidatos participam do estudo (por exemplo, profissionais, acadêmicos, consultores, estudantes), e as tarefas a ser executadas (por exemplo em pequena escala, grandes escala);

Não é necessário usar todos os elementos da abordagem PICOC, isto dependerá do objetivo da revisão sistemática. No quadro 8 pode se apreciar um exemplo onde se adapta a abordagem PICOC a fim de recolher provas sobre a previsão de manutenção de aplicações de *software*.

População	Aplicações de software
Intervenção	Métodos e técnicas para prever manutenção, indicadores de manutenção
Resultados	Exactidão dos métodos de previsão de manutenção / técnicas
Contexto	O contexto, neste caso vai ser abrangente. Este estudo irá abranger as universidades, bem como a indústria de software e todos os tipos de intervenientes (desenvolvedores, mantenedores, testadores, gerentes de projeto, estudantes e investigadores, etc). Ele também irá abranger todos os tipos de estudos empíricos, incluindo observações, entrevistas, questionários, pesquisas, experimentação formal e estudos de caso, etc.

Tabela 8 – Adaptação da abordagem PICOC em uma revisão sistemática. Adaptado de Riaz, Mendes e Tempero (2008).

3.1.2 Etapa 2: Execução da revisão

Uma pesquisa abrangente e imparcial é uma das diferenças fundamentais entre uma revisão narrativa tradicional e uma revisão sistemática. Uma busca sistemática começa com a identificação de palavras-chave e termos de pesquisa, que são construídos a partir do estudo exploratório, a literatura e discussões dentro da equipe de revisão. A estratégia da busca deve ser relatada com os detalhes suficientes para garantir que a pesquisa possa ser replicada. Pesquisas não deve ser realizadas apenas em revistas publicadas e bases de dados bibliográficos, mas também compreendem estudos inéditos, conferências, estudos da indústria, a Internet e até mesmo pedidos pessoais para os investigadores conhecidos (TRANFIELD; DENYER; SMART, 2003, p. 215). Para reduzir o erro humano e o viés, revisões sistemáticas utilizam formulários de extração de dados. Estas costumam conter informações gerais (título, autor, detalhes da publicação), as características do estudo e informações específicas (dados e métodos) e notas sobre temas emergentes juntamente com detalhes de síntese. (TRANFIELD; DENYER; SMART, 2003, p. 217). Síntese de pesquisa é o termo usado para uma família de métodos de síntese, integração e, se possível, acumulação dos resultados dos diferentes estudos sobre um tópico ou questão de pesquisa. A forma mais simples e mais conhecida de síntese é uma revisão narrativa que tenta identificar o que tem sido escrito sobre um assunto ou tema. Meta-análise é uma abordagem alternativa para a síntese, o que permite compartilhar os dados de estudos individuais para permitir um aumento do poder estatístico e uma estimativa mais precisa (TRANFIELD; DENYER; SMART, 2003, p. 217).

3.1.3 Etapa 3: Relatórios e divulgação

O pesquisador deve ser capaz de fornecer ampla descrição do campo, com exemplos específicos e justificando as suas conclusões. Os pesquisadores também precisam de um relatório dos resultados de uma “análise temática”, indicando se os resultados foram obtidos ou não através de uma abordagem agregadora ou interpretativa, destacando que é já conhecido a partir de formas de extração de dados das contribuições fundamentais (TRANFIELD; DENYER; SMART, 2003, p. 218).

3.2 VISUALIZAÇÃO DA INFORMAÇÃO

Nesta seção apresenta-se a revisão sistemática feita com o intuito de estabelecer o estado da arte da área de visualização da informação em relação aos temas de interesse desta tese. As perguntas que nortearam esta revisão são apresentadas na seqüência:

- **P1.** Como a área da visualização da informação tem se relacionado com as abordagens extração da informação e recuperação do conhecimento?
- **P2.** Que tipos de estruturas de visualização têm sido usados nos sistemas de visualização?
- **P3.** Como podem ser avaliados os sistemas de visualização de informação?

Para responder a pergunta **P1**. *Como a área da visualização da informação tem se relacionado com as abordagens extração da informação e recuperação do conhecimento?*, foram executadas as seguintes consultas nos indexadores *Scopus*, *ScienceDirect*, *IEEEExplore*, *ACM Digital Library*.

```
(*) Parâmetros utilizados no indexador Scopus
TITLE("information visualization" OR visualization) AND
TITLE("information extraction" OR "information retrieval")
```

```
(*) Parâmetros utilizados no indexador ScienceDirect
TITLE("information visualization" OR visualization) AND
TITLE("information extraction" OR "information retrieval")
```

```
(*) Parâmetros utilizados no indexador IEEEExplore
(("Document Title":"information visualization" OR
"Document Title":visualization)
AND ("Document Title":"information extraction" OR
"Document Title":"information retrieval"))
```

```
(*) Parâmetros utilizados no indexador ACM Digital Library
((Title:"information visualization" OR Title:visualization)
AND
(Title:"information extraction" OR
Title:"information retrieval"))
```

As buscas foram executadas no dia 4 de janeiro do ano 2011 retornando 35 estudos no indexador *Scopus*, 1 estudo no indexador *ScienceDirect*, 8 estudo no indexador *IEEEExplore* e 4 estudos no indexador *ACM Digital Library*. Foram selecionados somente os estudos onde as abordagens de extração de informação ou recuperação de informação foram consideradas como parte de sistemas de visualização de informação, desconsiderando resumos e introduções a anais de congressos. Também foram desconsiderados os estudos onde não se explicitava a forma em que eram usadas as visualizações. Após a aplicação dos critérios de exclusão foram analisados 14 estudos. A quantidade de estudos selecionados por cada indexador pode ser visualizada na tabela 9. Como é comum nas revisões sistemáticas, alguns estudos foram encontrados em mais de um indexador.

Indexador	Documentos retornados	Documentos selecionados
<i>Scopus</i>	35	15
<i>ScienceDirect</i>	1	0
<i>IEEEExplore</i>	8	2
<i>ACM Digital Library</i>	4	2

Tabela 9 – Resultados das buscas para a revisão sistemática na área da Visualização da Informação - pergunta P1.

A recuperação da informação pode-se definir como o campo que se preocupa pelos processos envolvidos na representação, armazenamento, busca e localização de informação que são relevantes aos requerimentos de informação desejados pelos usuários. De forma geral o problema consiste em procurar a informação em forma de texto e outros meios que satisfaçam otimamente o estado de incerteza do usuário, tendo em consideração que algumas informações serão mais relevantes que outras dependendo dos requerimentos do usuário (INGWERSEN, 1992, p. 49-50). Por outro lado, a extração da informação pode-se definir como um conjunto de técnicas que, a partir de um conjunto de textos de entrada (geralmente escrito em linguagem natural), permite a extração da informação e sua representação em uma forma estruturada e não ambígua (CUNNINGHAM, 1997, p. 1). Neste sentido, informação estruturada é criada a partir de textos não estruturados (FELDMAN et al., 2002), ou dito de outra forma, o processo de extração de informação pode ser visto como um conjunto de atividades que visam popular uma fonte de informação estruturada (por exemplo um banco de dados re-

lacional) a partir de uma não estruturada ou a partir de texto livre (GAIZAUSKAS; WILKS, 1998, p. 70).

Após analisar os estudos, verificou-se que, no âmbito da extração da informação e a recuperação da informação, os métodos e técnicas da visualização da informação foram usados para atingir quatro grandes objetivos: i) a exploração da informação, ii) a visualização da informação extraída, iii) a visualização da informação recuperada e iv) a visualização do histórico das buscas. A classificação destes estudos em cada grupo pode ser visualizada na tabela 10.

Em relação à exploração da informação, a visualização da informação ajuda na compreensão dos dados armazenados nos repositórios digitais, sobretudo quando estes repositórios possuem grandes quantidades de informação. Esta exploração da informação se dá de forma interativa em um processo onde o usuário pode alterar os parâmetros de visualização.

Em relação à visualização da informação extraída, a visualização da informação fornece os métodos necessários aos sistemas de extração da informação para poder visualizar as entidades e as relações extraídas a partir dos documentos. Estas visualizações são úteis quando é preciso reconhecer rapidamente as relações que existem entre as entidades extraídas. De forma similar, a visualização da informação ajuda aos sistemas de recuperação da informação na visualização da informação recuperada.

Objetivo	Autores
exploração da informação	Song (2009)
visualizar informação extraída	Seifert I. (2009), Christian, Patrick e Christophe (2006), Fahmi I. (2007), Chen A.A. (2005)
visualizar informação recuperada	Yoshinaga e Nobuhara (2010), Su, Feng e Ma (2010), Peltonen et al. (2010), Venna J. (2010), Castro P. (2009), Peltonen (2009), Zhang e Deng (2008), Rydberg-Cox J.A. (2004), Zhang J. (2002)
visualizar o histórico das buscas	Leide J.E. (2007)

Tabela 10 – Classificação das buscas para a revisão sistemática na área da Visualização da Informação - pergunta P1.

Para responder a pergunta **P2**. *Que tipos de estruturas de visualização têm sido usados nos sistemas de visualização?*, foram analisados os mesmos estudos selecionados na pergunta 1.

Em relação às estruturas de visualização usadas nos sistemas de visualização, verificou-se que o agrupamento da informação recuperada e/ou extraída em forma de *clusters* foi a estrutura mais usada nos estudos revisados (ZHANG J., 2002; RYDBERG-COX J.A., 2004; FAHMI I., 2007; PELTONEN, 2009; YOSHINAGA; NOBUHARA, 2010; PELTONEN et al., 2010; ZHANG; DENG, 2008; VENNA J., 2010). Também foram usadas as metáforas visuais (como a metáfora do magnetismo em Castro P. (2009)), os arvores (SEIFERT I., 2009), *treemaps* (SONG, 2009), mapas cartesianos e (SU; FENG; MA, 2010). A única representação do mundo real foi apresentada por Christian, Patrick e Christophe (2006) na qual documentos culturais eram apresentados usando mapas geográficos.

Por outro lado, em relação ao domínio de aplicação, verificou-se que a maioria dos estudos analisados foram propostos para trabalhar com documentos de forma geral (ZHANG J., 2002; RYDBERG-COX J.A., 2004; CHRISTIAN; PATRICK; CHRISTOPHE, 2006; FAHMI I., 2007; LEIDE J.E., 2007; VENNA J., 2010; SU; FENG; MA, 2010), mas também muitos trabalhos foram desenvolvidos para que sejam aplicados a domínios específicos como a visualização de artigos científicos (SEIFERT I., 2009; PELTONEN, 2009), páginas Web (YOSHINAGA; NOBUHARA, 2010; ZHANG; DENG, 2008), correios eletrônicos (CASTRO P., 2009), foros (SONG, 2009), genoma humano (PELTONEN et al., 2010) e a medicina (CHEN A.A., 2005)

Para responder a pergunta **P3**. *Como podem ser avaliados os sistemas de visualização de informação?*, foram executadas as seguintes consultas nos indexadores *Scopus*, *ScienceDirect*, *IEEEExplore*, *ACM Digital Library*.

```
(*) Parâmetros utilizados no indexador Scopus
(TITLE("information visualization") OR
TITLE("knowledge visualization")) AND
TITLE(measure OR measurement OR evaluation OR appraisal)
```

```
(*) Parâmetros utilizados no indexador ScienceDirect
(TITLE("information visualization") OR
TITLE("knowledge visualization")) AND
TITLE(measure OR measurement OR evaluation OR appraisal)
```

```
(*) Parâmetros utilizados no indexador IEEEExplore
(("Document Title": "information visualization") AND
```

```
("Document Title":measure OR
"Document Title":measurement OR
"Document Title":evaluation OR
"Document Title":appraisal))
```

```
(*) Parâmetros utilizados no indexador ACM Digital Library
((Title:"information visualization") AND
(Title:measure OR Title:measurement OR
Title:evaluation OR Title:appraisal))
```

As buscas foram executadas no dia 9 de janeiro do ano 2011 retornando 29 estudos no indexador *Scopus*, 2 estudos no indexador *ScienceDirect*, 6 estudo no indexador *IEEEExplore* e 10 estudos no indexador *ACM Digital Library*. Foram selecionados somente os estudos onde se propõem métodos e/ou técnicas para avaliar sistemas visualização, desconsiderando resumos e introduções a anais de congressos. Após a aplicação dos critérios de exclusão foram analisados 15 estudos. A quantidade de estudos selecionados por cada indexador pode ser visualizada na tabela 11. Como é comum nas revisões sistemáticas, alguns estudos foram encontrados em mais de um indexador.

Indexador	Documentos retornados	Documentos selecionados
<i>Scopus</i>	29	13
<i>ScienceDirect</i>	2	2
<i>IEEEExplore</i>	6	3
<i>ACM Digital Library</i>	10	7

Tabela 11 – Resultados das buscas para a revisão sistemática na área da Visualização da Informação - pergunta P3.

Um dos papéis dos pesquisadores da visualização da informação é fornecer evidências convincentes da utilidade dos sistemas de visualização, que é difícil para qualquer nova tecnologia, mas que apresenta desafios específicos para a visualização da informação. A usabilidade das ferramentas de visualização da informação pode ser medida em um laboratório, no entanto, para ser convincente, a utilidade deve ser demonstrada em um cenário real, ou seja, em um domínio de aplicação determinado e com um conjunto de usuários (PLAISANT, 2004), mas algumas características de visualização da informação tornam a sua avaliação particularmente desafiadora: os usuários muitas vezes preci-

sam olhar para os mesmos dados a partir de perspectivas diferentes e durante um longo tempo; eles também podem ser capazes de formular e responder perguntas que não tinham previsto antes de olhar para a visualização; e finalmente, apesar das descobertas terem um impacto enorme ocorrem muito raramente, ou até podem não ocorrer (PLAISANT, 2004; SHNEIDERMAN; PLAISANT, 2006).

Diferentes métodos podem ser utilizados para a avaliação de sistemas de visualização. Os mais comumente adotados são os métodos baseados no usuário e os métodos de inspeção. Os métodos baseados nos usuários consistem principalmente de testes com usuários, onde propriedades de usabilidade são avaliadas pela observação de como o sistema é utilizado por alguns representantes de usuários reais (DIX; FINLAY; ABOWD, 2003). A avaliação baseada em usuários atualmente oferece a mais completa forma de avaliação, porque avalia a usabilidade por meio de amostras em usuários reais. No entanto, esta técnica tem uma série de inconvenientes, como o sistema geralmente é avaliado com os usuários no campo apenas no final do seu desenvolvimento, os resultados da avaliação chegam tarde demais para serem considerados pelos desenvolvedores durante a fase de desenvolvimento. Desta forma, não é possível evitar erros graves ou tempo de reimplementação (ARDITO et al., 2006).

Os métodos de inspeção envolvem apenas peritos avaliadores, que controlam a aplicação e fornecem julgamentos baseados em seus conhecimentos (NIELSEN; MACK, 1994). Com relação à avaliação baseada no usuário, os métodos de inspeção de usabilidade são mais subjetivos, tendo uma forte dependência das habilidades do inspetor. Sua principal vantagem, no entanto, é a redução de custos: eles não precisam os usuários, e não requer nenhum equipamento especial, nem instalações de laboratório (NIELSON, 1993). Além disso, os especialistas podem detectar uma ampla gama de problemas e possíveis falhas de um sistema complexo em uma quantidade limitada de tempo. Dentre os métodos de inspeção, o mais comumente utilizado é a avaliação heurística (ARDITO et al., 2006). Trata-se de um pequeno conjunto de peritos inspecionando o sistema e avaliando a interface com uma lista de princípios de usabilidade reconhecidos: a heurística (ZUK et al., 2006).

Os estudos analisados podem ser visualizados na tabela 12. Na seqüência, apresenta-se uma breve descrição da revisão feita nos estudos analisados.

Segundo Forsell (2010) há uma variedade de objetivos para perseguir uma avaliação. Entre elas podemos mencionar i) avaliar os pontos fortes e fracos de diferentes técnicas, ii) procurar *insight* para deter-

minar por que uma técnica em particular oferece bom desempenho, iii) demonstrar que uma nova técnica é útil em um sentido prático de acordo com alguns critérios objetivos, iv) demonstrar que uma técnica nova é melhor do que uma já existente de acordo com alguns critérios objetivos, e v) investigar se os princípios teóricos de outras disciplinas se aplicam sob certas condições práticas. Por outro lado, Plaisant (2004) aponta quatro grandes áreas temáticas de avaliação: i) experimentos controlados comparando os elementos de design, ii) avaliação de usabilidade de uma ferramenta, iii) experimentos controlados comparando duas ou mais ferramentas, e iv) estudos de caso de ferramentas em cenários realistas.

Na pesquisa experimental, se investigam e comparam as respostas dos participantes em diferentes condições. Manipulando um (ou mais) fatores ou variáveis (variável independente) para investigar seu efeito sobre um (ou mais) outro fator (a variável dependente). Usando o paradigma quantitativo, Bresciani e Eppler (2009) apresentam um experimento no qual foram analisados os efeitos de visualização sobre o compartilhamento de conhecimento em grupos de trabalho. Os resultados estatísticos deste estudo mostram claramente que a visualização interativa tem um impacto estatisticamente significativo, objetivo e positivo nos resultados de compartilhamento de conhecimentos, ainda quando os sujeitos parecem não estar conscientes disso. Em particular, os grupos apoiados pela visualização alcançaram maior produtividade, maior qualidade dos resultados e maior ganho de conhecimento. Stasko et al. (2000) descreve uma avaliação de duas ferramentas de visualização de informações usadas para apresentar as hierarquias, especificamente estruturas, o diretório do computador e arquivo. Pillat, Valiati e Freitas (2005) apresentam um estudo experimental com base em testes com os usuários para avaliar duas técnicas de visualização de informações multidimensionais. Burkhard e Meier (2005) avalia seu modelo de visualização do conhecimento com usuarios, considerando as seguintes características: atenção, visão global, detalhes, metáfora, motivação e discussão. Chittaro e Combi (2001), da mesma forma, avalia qual representação de espaços temporais é a mais eficiente a partir do ponto de vista dos usuários.

A avaliação heurística é um método de avaliação conhecido e popular na área de Interação Humano-Computador. A avaliação heurística envolve um pequeno número de avaliadores inspecionando um sistema de acordo com heurísticas ou diretrizes que são relevantes para o sistema (ZUK et al., 2006). O método é fácil de aprender e aplicar, pode ser utilizado durante todas as fases de desenvolvimento e que requer de

pouco tempo e recursos. Além disso, ele pode ser modificado de várias maneiras para melhorar o resultado, por exemplo, incluindo especialistas de domínio ou usuários finais, incluindo considerações específicas do usuário, ou realizando tarefas específicas (FORSELL; JOHANSSON, 2010). Na linha da avaliação heurística dois estudos fazem uma meta-análise com o objetivo de aumentar entendimento do uso de diferentes conjuntos de heurísticas para avaliação de visualização de informações, Zuk et al. (2006) agrupam as heurísticas em três grupos e Forsell e Johansson (2010) em seis, estes grupos abrangem uma grande variedade de perspectivas sobre usabilidade relevantes para a visualização da informação. Winckler, Palanque e Freitas (2004) por seu lado, propõem modelos de tarefas como um elemento fundamental para a realização de avaliações de uma forma estruturada. Tarefas relativas à utilização de sistemas de visualização de informações devem ser consideradas em diferentes níveis de abstração. No nível mais alto (mais abstrato) encontra-se o objetivo do usuário e as suas intenções. Depois segue o nível abstrato que abrange um conjunto de subtarefas genéricas que são necessárias para cumprir a meta do usuário. O terceiro nível, o nível de interação, inclui a especificação da atividade realizada pelo usuário, o sistema ou ambos (tarefas interativas). O nível mais baixo, chamado de apresentação visual, inclui uma descrição concreta da atividade do sistema em termos de prestação da função utilizada para fornecer o usuário com informações sobre os dados.

O paradigma qualitativo também tem sido usado para avaliar os sistemas de visualização, com este tenta-se alcançar uma compreensão mais rica através de uma abordagem holística, considerando a interação entre os fatores que influenciam visualizações, seu desenvolvimento e seu uso. Usando este paradigma Valiati, Freitas e Pimenta (2008) baseado na observação participante e entrevistas, avaliam as de técnicas de visualização das informações multidimensional, com o objetivo de comparar os resultados de diferentes estudos de caso a respeito dos comportamento dos usuários, as tarefas executadas e o impacto dos problemas de usabilidade nas actividades dos usuários. Isenberg et al. (2008) introduzem *grounded evaluation* como um processo que tenta assegurar que a avaliação de uma ferramenta de visualização da informação está situada dentro do contexto da sua utilização. Kumar, Subramanian e Zhang (2008) propõem usar uma nova técnica chamada de abordagem NFR para avaliar qualitativamente as característica de "visuabilidade" das ferramentas de visualização de informação.

3.3 VISUALIZAÇÃO DO CONHECIMENTO

Nesta seção apresenta-se a revisão sistemática feita com o intuito de estabelecer o estado da arte da área de visualização do conhecimento em relação aos temas de interesse desta tese. As perguntas que nortearam esta revisão são apresentadas na seqüência:

- **P1.** Quais os framework existentes na área de visualização do conhecimento que suportam o desenvolvimento de aplicações baseadas em conhecimento?
- **P2.** Quais são os modelos de visualização do conhecimento propostos na literatura científica?

Para responder a pergunta **P1**. *Quais os framework existentes na área de visualização do conhecimento que suportam o desenvolvimento de aplicações baseadas em conhecimento?*, foram executadas as seguintes consultas nos indexadores *Scopus*, *ScienceDirect*, *IEEEExplore*, *ACM Digital Library*.

```
(*) Parâmetros utilizados no indexador Scopus
TITLE-ABS-KEY("knowledge visualization") AND
TITLE-ABS-KEY(framework)
```

```
(*) Parâmetros utilizados no indexador ScienceDirect
TITLE-ABSTR-KEY("knowledge visualization") AND
TITLE-ABSTR-KEY(framework)
```

```
(*) Parâmetros utilizados no indexador IEEEExplore
("Document Title":"knowledge visualization" AND
 "Document Title":framework)
OR ("Abstract":"knowledge visualization" AND
 "Abstract":framework)
OR ("Author Keywords":"knowledge visualization" AND
 "Author Keywords":framework)
```

```
(*) Parâmetros utilizados no indexador ACM Digital Library
((Title:"knowledge visualization" AND Title:model) OR
 (Abstract:"knowledge visualization" AND
  Abstract:framework) OR
 Owner:ACM(Keywords:"knowledge visualization" AND
 Keywords:framework))
```

As buscas foram executadas no dia 16 de janeiro do ano 2011 retornando 26 estudos no indexador *Scopus*, 0 estudos no indexador *ScienceDirect*, 7 estudo no indexador *IEEEExplore* e 9 estudos no indexador *ACM Digital Library*. Foram selecionados somente os estudos onde se propõem *frameworks* para desenvolver aplicações baseadas em conhecimento, desconsiderando resumos e introduções a anais de congressos. Após a aplicação dos critérios de exclusão foram analisados 8 estudos. A quantidade de estudos selecionados por cada indexador pode ser visualizada na tabela 13. Como é comum nas revisões sistemáticas, alguns estudos foram encontrados em mais de um indexador. Na seqüência, apresenta-se uma breve descrição de cada estudo.

- Mohd, Embong e Zain (2010) propõem um *framework* para a visualização do conhecimento através da adoção de gráficos baseado em técnica de visualização e incorporando conceito *dashboard* para instituições de ensino superior, enfatizando a interação humano-computador. O *framework* consiste no modelo de *dashboard*, a arquitetura do sistema e o protótipo aplicado no ambiente de ensino superior.
- Gu, Ahmad e Sumner (2010) propõem um *framework* que usando processamento de linguagem natural e algoritmos baseados em grafos, permita visualizar: o conhecimento previo dos usuários, o conhecimento do domínio, novos conceitos encontrados visando revelar as relações semânticas entre eles. Assim, ajuda-se aos usuário a resolver suas incertezas em relação as novas idéias e conceitos que se encontram no decorrer do processo de aprendizagem, a fim de integrar novos conhecimentos com os seus conhecimentos prévios.
- Luo et al. (2007) e Luo H. (2007) propõem um *framework* que possibilita a mineração intuitiva e exploração de grandes bases de dados de notícias de vídeo através usando métodos da área da visualização do conhecimento.
- No trabalho de Burkhard (2004), Burkhard (2005c), Burkhard et al. (2005), Burkhard e Meier (2005) apresenta as sinergias entre as pesquisas na visualização de informações e visualização do conhecimento a partir de uma perspectiva da gestão do conhecimento e da ciência da comunicação. Apresenta um *framework* teórico e um modelo para o novo campo da visualização do conhecimento.

Para responder a pergunta **P2**. *Quais são os modelos de visualização do conhecimento propostos na literatura científica?*, foram

executadas as seguintes consultas nos indexadores *Scopus*, *ScienceDirect*, *IEEEExplore*, *ACM Digital Library*.

(*) Parâmetros utilizados no indexador Scopus
TITLE-ABS-KEY("knowledge visualization") AND
TITLE-ABS-KEY(model)

(*) Parâmetros utilizados no indexador ScienceDirect
TITLE-ABSTR-KEY("knowledge visualization") AND
TITLE-ABSTR-KEY(model)

(*) Parâmetros utilizados no indexador IEEEExplore
("Document Title":"knowledge visualization" AND
"Document Title":model)
OR ("Abstract":"knowledge visualization" AND
"Abstract":model)
OR ("Author Keywords":"knowledge visualization" AND
"Author Keywords":model)

(*) Parâmetros utilizados no indexador ACM Digital Library
((Title:"knowledge visualization" AND Title:model) OR
(Abstract:"knowledge visualization" AND Abstract:model) OR
Owner:ACM(Keywords:"knowledge visualization"
AND Keywords:model))

As buscas foram executadas no dia 20 de janeiro do ano 2011 retornando 54 estudos no indexador *Scopus*, 1 estudos no indexador *ScienceDirect*, 18 estudo no indexador *IEEEExplore* e 3 estudos no indexador *ACM Digital Library*. Foram selecionados somente os estudos onde se propõem modelos para a visualização do conhecimento, desconsiderando resumos e introduções a anais de congressos. Após a aplicação dos critérios de exclusão foram analisados 9 estudos. A quantidade de estudos selecionados por cada indexador pode ser visualizada na tabela 14. Como é comum nas revisões sistemáticas, alguns estudos foram encontrados em mais de um indexador. Na seqüência, apresenta-se um resumo de cada estudo.

- Burkhard (2005c) O objetivo deste modelo é a visualização do conhecimento genérico baseado em visualizações complementares, onde as representações visuais a usar dependerão das competências do grupo alvo. O modelo descreve processos iterativos inter e intrapessoais. O processo começa com um remetente que

pretende transferir alguns dos seus conhecimentos para um destinatário.

- Han e Cercone (2000a); Han e Cercone (2000b) Este modelo tem por objetivo a visualização das fases do processo de descoberta do conhecimento e mineração de dados. O modelo usa técnicas de visualização para: desenhar os dados originais e obter *insights* sobre a distribuição dos dados, intuitivamente limitar o domínio de dados através da interação, descobrir as regras a partir dos dados reduzidos e, finalmente, visualizar o conhecimento resultante. A visualização das regras é feita usando polígonos.
- Novak e Wurst (2005) A idéia principal deste modelo é o uso de artefatos de conhecimento dinâmicos que apoie aos usuário no acesso aos espaços de informação da comunidade, de uma maneira, que permite a descoberta de relações do conhecimento nas diferentes comunidades. Neste modelo se usa a metáfora do mapa do conhecimento como um veículo para descrever a idéia de uma estruturação visual da informação.
- Luo et al. (2007) O modelo proposto aqui permite a mineração intuitiva e a exploração de bases de dados de vídeos de notícias em grande escala através da visualização do conhecimento. Para apoiar a extração e visualização de conhecimento, um algoritmo de análise estatística dos vídeos é proposto para extrair a semântica dos relatórios dos vídeos. A visualização dos resultados das buscas se apresenta usando metáforas visuais.
- Li e Lu (2010) O objetivo deste modelo é ajudar os usuários a que possam visualizar e adquirir conhecimento a partir de fontes de informação distribuídas e heterogêneas, usando para este propósito mapas de tópicos inteligentes. O modelo inclui mecanismos que permitem a organização do conhecimento, a recomendação do conhecimento e a visualização do conhecimento.
- Mohd, Embong e Zain (2010) Neste modelo se propõe um modelo para a visualização do conhecimento adotando a técnica de visualização de gráfico baseado em conceitos para o domínio das instituições de ensino superior. A visualização permite a análise das características globais dos alunos, professores e indivíduos. A visualização do conhecimento é reforçada através do conceito de *dashboard* onde se prevê padrões significativo de conhecimentos.

- Liang e Huang (2010) Propõem um modelo baseado em camadas que tem por objetivo a visualização de elementos destacados. Os resultados da visualização são apresentados usando metáforas visuais.
- Melgar, Beppler e Pacheco (2010a) Propõe-se uma arquitetura para a visualização do conhecimento baseada em imagens semânticas no domínio da anatomia. Este trabalho é um dos resultados desta pesquisa.

3.4 WEB SEMÂNTICA E ANOTAÇÃO SEMÂNTICA DE IMAGENS

A pergunta que norteou esta revisão sistemática foi *Como tem sido feitas as anotações de imagens no contexto da Web Semântica?*. Para responder a pergunta, foram executadas as seguintes consultas nos indexadores *Scopus*, *ScienceDirect*, *IEEEExplore*, *ACM Digital Library*.

(*) Parâmetros utilizados no indexador Scopus

```
(TITLE-ABS-KEY("semantic web") AND
TITLE-ABS-KEY("image annotation"))
```

(*) Parâmetros utilizados no indexador ScienceDirect

```
(TITLE-ABSTR-KEY("semantic web") AND
TITLE-ABSTR-KEY("image annotation"))
```

(*) Parâmetros utilizados no indexador IEEEExplore

```
("Document Title":"semantic web" AND
"Document Title":"image annotation") OR
("Abstract":"semantic web" AND
"Abstract":"image annotation") OR
("Author Keywords":"semantic web" AND
"Author Keywords":"image annotation")
```

(*) Parâmetros utilizados no indexador ACM Digital Library

```
((Title:"semantic web" AND Title:"image annotation") OR
(Abstract:"semantic web" AND Abstract:"image annotation")
OR Owner:ACM(Keywords:"semantic web" AND
Keywords:"image annotation"))
```

As buscas foram executadas no dia 5 de fevereiro do ano 2011 retornando 50 estudos no indexador *Scopus*, 2 estudos no indexador

ScienceDirect, 2 estudo no indexador *IEEEExplore* e 1 estudos no indexador *ACM Digital Library*. Foram selecionados somente os estudos onde se explica claramente como é feita a anotação das imagens. Após a aplicação dos critérios de exclusão foram analisados 29 estudos. A quantidade de estudos selecionados por cada indexador pode ser visualizada na tabela 15. Como é comum nas revisões sistemáticas, alguns estudos foram encontrados em mais de um indexador. Na seqüência, apresenta-se uma análise desta revisão.

Dos estudos selecionados, estes se dividiam em três grupos com três objetivos diferentes. O primeiro deles abarcava o estudo das anotações automática das imagens (BARRAT; TABBONE, 2010; DING et al., 2009; DING; XU, 2010; GUO; LIAO, 2009; JIN Y., 2010; REN, 2010; LU et al., 2008; XU H.-T., 2010; ZHENG, 2010). O segundo grupo abarcava o estudo de anotações de imagens no contexto das redes sociais com o intuito de sugerir anotações para as imagens dos usuários das redes (ELAHI; KARLSEN; AKSELSEN, 2009; ELAHI N., 2010; ELAHI; KARLSEN; YOUNAS, 2011). O terceiro grupo se refere a anotação não automática de imagens que se detalha na seqüência.

Os metadados associados às imagens podem ser classificados como: i) metadados com conteúdo independente, quando estes se encontram relacionados à imagem, mas não a descrevem (*i.e.*, autores, data de criação, etc.); ii) metadados com conteúdo dependente, quando se referem a características de baixo nível e/ou nível intermediário (*i.e.*, cor, textura, etc.); iii) metadados com conteúdo descritivo, quando se referem a conteúdo semântico (*i.e.*, relações das entidades da imagem com entidades do mundo real) (HANBURY, 2008).

Os metadados com conteúdo descritivo podem ser especificados usando um ou mais dos seguintes enfoques: i) descrições usando texto livre, quando nenhuma estrutura predefinida para a anotação é proporcionada; ii) classificação usando palavras-chave, quando palavras-chave selecionadas de forma arbitrária ou a partir de vocabulários controlados são usadas para descrever as imagens; iii) classificação baseada em ontologias, quando conceitos ou instância de uma ontologia são usados no processo de anotação (MÜLLER et al., 2004; HANBURY, 2008).

Por outro lado, os metadados com conteúdo descritivo podem ser proporcionados a dois níveis de especificidade: i) conteúdo descritivo associado à imagem completa (SCHREIBER et al., 2001, 2002; HYVÖNEN; STYRMAN, 2002; HOLLINK et al., 2003; OSMAN et al., 2007) e ii) segmentação da imagem com vinculação do conteúdo descritivo em cada região da imagem (CHRONAKI; ZABULIS; ORPHANOUDAKIS, 1997; BOTTONI et al., 2006; HALASCHEK-WIENER et al., 2005b, 2005, 2005a, 2006;

MÖLLER; SINTEK, 2007; RUBIN et al., 2008a, 2008b, 2009; HSU et al., 2009; SONNTAG; MÖLLER, 2010).

Os sistemas para anotar imagens usando metadados com conteúdo descritivo baseados em ontologias, geralmente usam dois tipos de ontologias, uma para definir o esquema de anotação e outra para definir os conceitos do domínio. Isto permite que os esquemas de anotação sejam definidos de forma independentes ao domínio do conhecimento. A forma de definir a ontologia de anotação depende dos requerimentos da aplicação. Por exemplo, Schreiber et al. (2001) usaram uma estrutura baseada em quatro elementos “agente - ação - objeto - cenário” para anotar imagens de macacos. Em Schreiber et al. (2002) um conjunto estruturado de propriedades de objetos foi usado para anotar imagens de móveis antigos. De forma similar em Hollink et al. (2003) foi usado os elementos “agente - ação - objeto - receptor” para anotar uma coleção de imagens de arte. Osman et al. (2007) usaram a estrutura “ator - ação - objeto” para anotar uma coleção de imagens no domínio dos esportes. Em algumas casos a associações de instâncias da ontologia a imagem é suficiente (HYVÖNEN; STYRMAN, 2002).

Autor	Paradigma	Breve descrição
Plaisant (2004)		Informa as dificuldades que existem ao avaliar sistemas de visualização
Forsell e Johansson (2010)	heurísticas	Foca-se na avaliação de sistemas de visualização por meio de heurísticas, é uma guia de aplicação
Zuk et al. (2006)	heurísticas	Foca-se na avaliação de sistemas de visualização por meio de heurísticas, é uma guia de aplicação
Winckler, Palanque e Freitas (2004)	heurísticas	Apresenta heurísticas baseada em modelo de tarefas
Amar e Stasko (2004)	heurísticas	Apresenta heurísticas baseada em tarefas de conhecimento
Valiati, Freitas e Pimenta (2008)	qualitativo	Compara sistemas visualização de informação multidimensional
Isenberg et al. (2008)	qualitativo	Verifica que a avaliação dos sistemas se situe no contexto da utilização do usuário
Kumar, Subramanian e Zhang (2008)	qualitativo	Avalia as características dos sistemas de visualização de informação
Forsell (2010)	quantitativo	Foca-se na pesquisa experimental quantitativa
Bresciani e Eppler (2009)	quantitativo	Verifica o uso das visualizações para o compartilhamento de conhecimento em grupos de trabalho
Pillat, Valiati e Freitas (2005)	quantitativo	Avalia duas técnicas de visualização da informação multidimensional
Burkhard e Meier (2005)	quantitativo	Avalia uma aplicação de visualização do conhecimento
Stasko et al. (2000)	quantitativo	Compara dos métodos de visualização
Chen e Czerwinski (2000)	quantitativo	Foca-se na pesquisa experimental quantitativa
Chittaro e Combi (2001)	quantitativo	Avalia qual representação de espaços temporais é a melhor

Tabela 12 – Estudos analisados para a revisão sistemática na área da Visualização da Informação - pergunta P3.

Indexador	Documentos retornados	Documentos selecionados
<i>Scopus</i>	26	8
<i>ScienceDirect</i>	0	0
<i>IEEEExplore</i>	7	4
<i>ACM Digital Library</i>	9	4

Tabela 13 – Resultados das buscas para a revisão sistemática na área da Visualização do Conhecimento - Pergunta 1.

Indexador	Documentos retornados	Documentos selecionados
<i>Scopus</i>	54	9
<i>ScienceDirect</i>	1	0
<i>IEEEExplore</i>	18	1
<i>ACM Digital Library</i>	3	0

Tabela 14 – Resultados das buscas para a revisão sistemática na área da Visualização do Conhecimento - Pergunta VC2.

Indexador	Documentos retornados	Documentos selecionados
<i>Scopus</i>	50	29
<i>ScienceDirect</i>	2	0
<i>IEEEExplore</i>	2	1
<i>ACM Digital Library</i>	1	1

Tabela 15 – Resultados das buscas para a revisão sistemática na área da anotação de imagens.

4 MODELO PROPOSTO

4.1 INTRODUÇÃO

Atualmente nos encontramos vivendo na era do conhecimento, onde as sociedades são construídas sobre a sinergia de indivíduos, grupos, organizações, redes sociais e comunidades que exploram formas eficazes de disseminação do conhecimento e fluxos de aprendizagem, onde a maior parte do valor adicionado na produção descansa sobre o conhecimento incorporado no produto final, onde o objetivo é uma sociedade com acesso ao conhecimento e aprendizagem para todos.

Esta era é caracterizada, por um lado, por uma explosão sem precedentes de informações que são digitalizadas e disponibilizadas para grandes audiências através de ambiente livres e *on-line*, e por outro lado, por uma sociedade cada vez mais exigente e ansiosa para aprender e saber aproveitar os benefícios destes conhecimentos. Uma consequência direta é que muita informação está sendo disponibilizada e as pessoas se tornam menos capazes de tomar decisões, pois não sabem a quem prestar atenção. Existe, portanto, uma grande necessidade de gerenciar esta grande quantidade de informações digitais e transformá-las em conhecimento utilizável.

Neste contexto, propõe-se um modelo para a visualização do conhecimento que está embasado nas áreas previamente pesquisadas. A visualização da informação auxilia a entender como os seres humanos processam as representações visuais. A visualização é um processo onde o usuário interage diretamente controlando os parâmetros de visualização com o objetivo de compreender o conjunto de dados. A metodologia CommonKADS, através de seu modelo de contexto, permitiu identificar os problemas e oportunidades no contexto da sociedade do conhecimento e ajudou a reconhecer quais os processos intensivos em conhecimentos que o modelo deve considerar. A visualização do conhecimento forneceu o *framework* que norteou o trabalho. Usando este *framework* foi possível dar resposta a questão: “Por que e como visualizar o conhecimento?”. Ao tentar responder esta pergunta é que surgem as imagens semânticas como meio para visualizar e recuperar conhecimento a partir dos repositórios digitais.

O modelo utiliza também mapeamentos semânticos que visam suportar a integração semântica tanto das imagens quanto dos repositórios de documentos. Estes mapeamentos semânticos permitem que o modelo “entenda” quais os conceitos representados nas imagens e nos

documentos, além disso, possibilita as consultas baseadas em conceitos como alternativa à recuperação de informação textual.

Este capítulo está organizado da seguinte forma após a introdução, se apresenta a aplicação da metodologia CommonKADS com o intuito de reconhecer os elementos do modelo. Posteriormente estes elementos são organizados no *framework* para visualização do conhecimento proposto por Burkhard. Finalmente é apresentado o modelo usando a visão sistemática de Bunge.

4.2 APLICAÇÃO DA METODOLOGIA COMMONKADS

A metodologia CommonKADS (SCHREIBER et al., 1999) possui um conjunto de modelos que permitem construir sistemas baseados em conhecimento. Dado que o objetivo deste trabalho não é criar um sistema baseado em conhecimento e sim um modelo, não foram usados todos os modelos desta metodologia. Foi usado apenas o modelo da organização visando caracterizar e estruturar a sociedade do conhecimento no contexto do problema formulado nesta pesquisa. Na figura 23 se pode apreciar a visão geral deste modelo. A construção deste modelo é feita usando planilhas (OM-1, OM-2, OM3 e OM-4 na figura 23) que permitem identificar os principais problemas e oportunidades assim como as relações das tarefas intensivas em conhecimento com os agentes envolvidos nos processos intensivos em conhecimento.

A primeira parte do modelo da organização foca-se nos problemas e oportunidades vistos desde um amplo contexto organizacional. O contexto organizacional abarca amplas categorias tais como a missão da organização, seus objetivos, estratégias, entre outros. Assume-se que este contexto é invariante, apesar disso, as oportunidades, os problemas e as propostas de soluções baseadas em conhecimento deverão sempre ser julgadas dentro desta ampla perspectiva organizacional. Para cumprir com este primeiro objetivo a metodologia CommonKADS disponibiliza uma planilha, chamada de OM-1, onde se descrevem os problemas e oportunidades, o contexto organizacional e as soluções propostas.

O objetivo principal da planilha OM-1 é descrever as organizações. Neste trabalho esta planilha foi adaptada para descrever as organizações no contexto da sociedade do conhecimento. Neste contexto, o conhecimento e o aprendizado são os principais objetivos da sociedade, mas esta sociedade encontra-se caracterizada por um crescimento exponencial da informação em formato digital, ocasionando um problema para os usuários que não podem ler e analisar toda a informação disponível.



Figura 23 – Modelo da organização em CommonKADS. Adaptado de Schreiber et al. (1999).

Neste cenário se vê como oportunidade a visão do mundo que compartilham grupos de usuários e as comunidades de prática, esta visão de mundo possibilita a utilização de imagens do mundo real como veículo para auxiliar nos processos de recuperação do conhecimento. O uso de imagens do mundo real possibilita a associação do conhecimento prévio dos usuários. A tabela 16 apresenta a planilha OM-1 usada nesta trabalho.

A segunda parte de modelo da organização se concentra sob aspectos mais específicos da organização. Aqui se cobrem aspectos como, por exemplo, a forma em que os processos do negócio são estruturados, os grupos de trabalho envolvidos nos processos de negócio, quais os recursos usados e assim por diante. Estes componentes do modelo da organização poderiam cambiar como resultados da introdução de sistemas baseados em conhecimento. Como ajuda para esta análise, a metodologia CommonKADS proporciona a planilha OM-2. Nesta planilha se detalha: a estrutura da organização incluindo departamentos, grupos, unidades, seções entre outros; os processos detalhando as suas tarefas internas; as pessoas ou agentes de *software* envolvidos; os recur-

Problemas	Dificuldade dos usuários para ler e analisar as informações disponíveis.
Oportunidades	Grupos e comunidades que compartilham uma mesma visão do mundo.
Contexto	Crescimento exponencial da informação em formato digital.
	Organizações no contexto da sociedade do conhecimento.
Soluções	Integrar as informações dos repositórios digitais e usar imagens do mundo real para ajudar no processo de recuperação do conhecimento.

Tabela 16 – Planilha OM-1 CommonKADS.

tos utilizados nos processos do negócio que poderiam ser de diferentes tipos como sistema de informação e outros recursos computacionais, equipamento, materiais, tecnologia, patentes, entre outros; o conhecimento; e cultura e poder dentro e fora da organização.

Na tabela 17 apresenta a planilha OM-2 usada neste trabalho. Nela pode se apreciar os principais processos suportados pelo modelo proposto: a visualização de imagens, a recuperação do conhecimento e a recuperação de documentos. Estes processos são executados por pessoas que compartilham uma mesma visão do mundo e portanto entendem as entidades representadas nas imagens do mundo real as quais formam parte desta visão do mundo. Usando estas imagens do mundo real se acessa aos repositórios de documentos digitais usando para este fim o conhecimento da visão do mundo e as entidades mencionadas nos documentos e nas imagens.

Na figura 24 se apresenta o diagrama de atividades, na linguagem UML, do processo Visualização da Imagem. Este processo tem como objetivo a busca de imagens que satisfaçam o critério de busca informado pelo usuário. A idéia deste processo vá além de retornar uma imagem, pois também consulta os documentos relacionados com as entidades representadas na imagem consultada pelo usuário. Desta forma o modelo pode informar ao usuário quantos documentos existem por cada entidade representada na imagem.

O processo Recuperação do Conhecimento, também tem sido detalhado usando o diagrama de atividades da linguagem UML. Este diagrama pode ser visualizado na figura 24, onde se pode observar como

Processos	P1. Visualização de imagens.
	P2. Recuperação do conhecimento.
	P3. Recuperação do documentos.
Pessoas	Usuários que compartilham uma mesma visão do mundo.
Recursos	R1. Repositórios digitais de documentos.
	R2. Imagens que representam entidades do mundo real.
Conhecimento	C1. Conhecimento relacionado ao domínio.
	C2. Entidades mencionadas nos documentos digitais.
	C3. Entidades representadas nas imagens.

Tabela 17 – Planilha OM-2 CommonKADS.

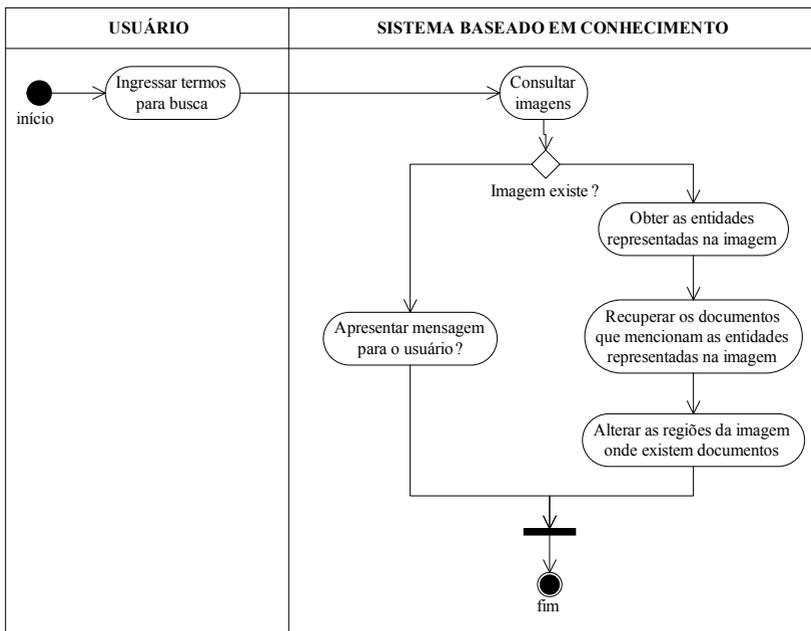


Figura 24 – Diagrama de seqüência do processo “visualização de imagens”.

se realiza a recuperação do conhecimento. O processo começa quando o usuário seleciona uma entidade representada na imagem, dado que se tem conhecimento explícito sobre a entidade, podem ser consultadas as relações existentes do conceito representado na entidade. O usuário pode usar estas relações para filtrar e refinar suas buscas. Por exemplo, se o usuário está visualizando uma imagem de um coração e tem selecionada a região da aorta, dado que se tem conhecimento do domínio, neste caso da anatomia, poderiam ser recuperadas relações tipo doenças relacionadas ou sintomas das doenças e usar as instâncias destas relações para refinar a busca.

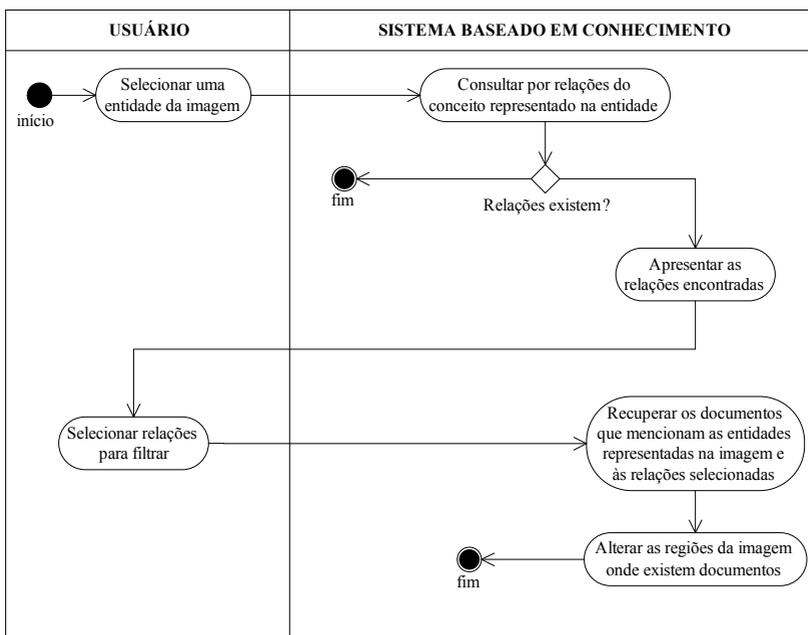


Figura 25 – Diagrama de seqüência do processo “recuperação do conhecimento”.

Os processos são também especificados com maior detalhe com a ajuda de uma planilha adicional. Nesta planilha os processos são detalhados em pequenas tarefas, pois, todo sistema previsto de conhecimento sempre realiza uma tarefa específica, e esta deve se encaixar adequadamente no processo como um todo. Frequentemente algumas adaptações dos processos são necessários requerindo algumas mudanças ou combinando-as com outras. Para ajudar nesta especificação, a meto-

dologia CommonKADS oferece a planilha OM-3 onde são especificados estes detalhes. Nesta planilha são especificados: o nome da tarefa, o agente que executa a tarefa, o local dentro da organização onde a tarefa é executada, o conhecimento usado, além de especificar se a tarefa é intensiva em conhecimento.

Na tabela 18 pode se observar a planilha OM-3 para o processo Visualização de Imagens, nela pode ser observar as tarefas definidas na figura 24 detalhando os agentes envolvidos e o conhecimento requerido para efetuar cada uma das tarefas. De forma similar na tabela 19 pode se observar a planilha OM-3 para o processo de Recuperação do Conhecimento, onde se detalham as tarefas definidas na figura 25.

O uso do modelo da organização da metodologia CommonKADS permitiu descrever os principais atores que serão representados no modelo da presente pesquisa, assim como determinar quais os conhecimentos necessários para executar, de forma satisfatória, cada uma das tarefas que compõem os processos intensivos em conhecimento identificados como possíveis soluções para o problema proposto neste trabalho. Este modelos conjuntamente com o *framework* para a visualização do conhecimento proposto por Burkhard (2005c), permitiram identificar os componentes do modelo proposto. A planilhas CommonKADS se complementa com el *framework* para a visualização do conhecimento, pois a metodologia CommonKADS está orientada para o desenvolvimento de sistemas baseados em conhecimento e o *framework* para a visualização para o desenvolvimento de sistemas de visualização do conhecimento. Na próxima seção se descreve o modelo no contexto da visualização do conhecimento.

4.3 MODELO NO CONTEXTO DA VISUALIZAÇÃO DO CONHECIMENTO

O modelo proposto apóia-se também no *framework* para visualização do conhecimento proposto por Burkhard (2005c). Neste *framework* cinco perguntas-chave são propostas para orientar a aplicação da visualização do conhecimento (ver tabela 20). Na seqüência se detalha as respostas às perguntas-chave que forneceram o quadro conceitual que norteou o desenvolvimento do modelo.

- **Que tipo de conhecimento precisa ser visualizado?** Conhecimento armazenado em repositório de documentos digitais. O conhecimento a ser visualizado é explicitado nos repositórios por meio de mapeamentos semânticos. Estes mapeamentos permitem

que a informação contida nos documentos possa ser entendida e processada automaticamente pelos computadores;

- **Por que o conhecimento deve ser visualizado?** Para facilitar a divulgação do conhecimento. Em decorrência aos avanços no processamento e gerenciamento eletrônico de documentos gerou-se uma acumulação grande de conhecimento nos repositórios digitais que tem excedido o que os usuários comuns podem perceber. Atualmente os usuários têm muitas dificuldades para lerem e analisarem as informações existentes devido ao pouco tempo disponível, à dificuldade de traduzir a suas necessidades de informação e a existência de múltiplas fontes de informação;
- **A quem está sendo destinado?** A um grupo de pessoas que compartilham uma mesma visão de mundo em um domínio específico. Ao compartilharem a mesma visão do mundo, os usuários pertencentes ao grupo em questão enxergam os conceitos da mesma forma e compartilham os mesmos modelos mentais.
- **Em que contexto deveria ser visualizado?** Em um entorno virtual, onde o conhecimento pode ser armazenado e codificado por meio de ontologias, taxonomias, tesouros e meta-tesouros, onde múltiplos mapeamentos podem ser feitos a um mesmo documento e a uma mesma representação visual;
- **Como pode o conhecimento ser visualizado?** Por meio de imagens que contêm representações visuais que *a priori* são conhecidas pelo grupo-alvo. As representações visuais são processadas rapidamente pelo cérebro humano. Além disso, ao serem conhecidas pelo grupo-alvo viabiliza a construção de modelos mentais que facilitam a associação da informação visual com conhecimento do mundo real.

4.4 DESCRIÇÃO DO MODELO

O modelo proposto foi concebido com o objetivo de facilitar a visualização do conhecimento usando como estruturas de apoio imagens semânticas. Na seqüência, descreve-se o modelo usando a visão de sistema de Bunge (2003). Segundo Bunge (2003) qualquer sistema concreto s pode ser modelado, em qualquer instante, pela quádrupla $\mu(s) = \langle C(s), E(s), S(s), M(s) \rangle$, onde:

- $C(s)$ = **Composição**: é a coleção de todas as partes de s ;
- $E(s)$ = **Ambiente**: é a coleção de itens que não pertencem a s e atuam ou sofrem a ação por algum ou todos os componentes de s ;
- $S(s)$ = **Estrutura**: coleção de relações, em particular ligações, entre os componentes de s ou entre esses e seu ambiente $E(s)$;
- $M(s)$ = **Mecanismos**: coleção de processos que fazem o sistema se comportar da maneira que se comporta;

Na tabela 21 se resumem os elementos do modelo. O modelo é composto pelos **Usuários**, **Imagens semânticas**, **Repositórios de documentos**, **Repositório de conhecimento** e o componente de **Visualização**. Os **Usuários** interagem no modelo por meio de interações com o componente de **Visualização**, é através destas interações que os **Usuários** selecionam os **Imagens Semânticas** e recuperam o conhecimento a partir dos **Repositórios de documentos**. O **Repositório de conhecimento** é usado pelo modelo para pode interpretar o conteúdo tanto dos **Imagens semânticas** quanto dos **Repositórios de documentos**.

O ambiente é caracterizado pela **Necessidade de informação** dos **Usuários**. É esta necessidade que move os **Usuários** a procurar as imagens para recuperar e visualizar conhecimento. É esta necessidade também que leva os **Usuários** a filtrar certo tipo de informação e a visualizar o conteúdo de certo conjunto de documentos.

A estrutura do modelo é composta pela **Visão de mundo** dos usuários e pelas **Anotações semânticas**. A **Visão de mundo** permite que os **Usuários** entendam os conceitos que estão representados nas **Imagens semânticas**. Por outro lado, as **Anotações semânticas** permitem que as **Imagens semânticas** e os **Repositórios de documentos** sejam integrados no modelo por meio do **Repositório de conhecimento**, possibilitando o seu processamento pelos computadores.

A **Visualização das imagens**, **Visualização do conhecimento**, **Refinamento**, e **Visualização dos documentos** compõem o conjunto de mecanismos do modelo. A **Visualização das imagens** é o mecanismo que permite que os usuários consultem e visualizem as **Imagens semânticas**; a **Visualização do conhecimento** é o mecanismo que consulta nos **Repositórios de documentos** e os conceitos mapeados nas **Imagens semânticas** e apresenta graficamente os resultados da busca; o **Refinamento** é o mecanismo que permite aos **Usuários** restringir o espaço da **Visualização do conhecimento**; e a

Visualização dos documentos é o mecanismo que permite recuperar um documento específico a partir do **Repositórios de documentos**.

4.4.1 COMPOSIÇÃO

4.4.1.1 Usuários

Os **Usuários** no modelo representam as pessoas que interagem com o componente de **visualização** motivados pela necessidade de informação que precisam satisfazer. Estes usuários estão caracterizados por compartilharem a mesma visão de mundo.

4.4.1.2 Imagens semânticas

As **Imagens semânticas** podem ser definidas como uma estrutura que permite representar graficamente os conceitos de um domínio, onde cada uma das partes que o compõem é especificada de forma explícita e formal, como por exemplo, em uma ontologia.

A imagem semântica é definida em três níveis: i) nível descritivo, ii) nível estrutural e iii) nível semântico. O **nível descritivo** visa identificar e descrever o elemento visual usado nas imagens semânticas, usando para isto atributos de identificação (*i.e.*, código da imagem), atributos físicos (*i.e.*, tipo de elemento visual, nome do arquivo físico, localização do arquivo físico, etc.) e atributos bibliográficos (*i.e.*, nome, autor/criador, data de criação, etc.). O **nível estrutural** fornece informações sobre a estrutura interna do imagem visando explicitar as regiões onde é possível identificar conceitos do domínio. Neste nível as regiões são etiquetadas com nomes escritos usando a linguagem natural. O **nível semântico** tem como objetivo enriquecer a imagem com conteúdo semântico (*i.e.*, classes ou instâncias definidas em uma ontologia). O conteúdo semântico é obtido a partir do repositório de conhecimento e vinculado às imagens por meio dos mapeamentos.

Na figura 26 apresenta-se um exemplo de uma imagem semântica usado para visualizar o conhecimento relacionado ao coração. O repositório de conhecimento foi composto pela ontologia FMA (Foundational Model of Anatomy), uma ontologia de referencia no domínio anatômico (ROSSE; MEJINO, 2003, 2008), e pelo metatesauro UMLS (Unified Medical Language System), um repositório que integra vários vocabulários biomédicos (BODENREIDER, 2004). Como elemento visual

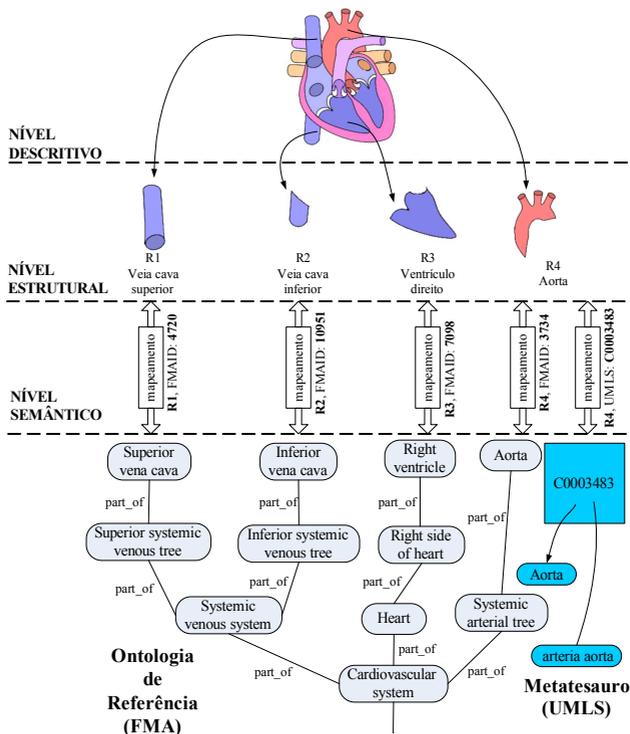


Figura 26 – Estrutura das imagens semânticas.

foi utilizado uma imagem em duas dimensões onde quatro regiões foram explicitadas: a veia cava superior, a veia cava inferior, o ventrículo direito e a aorta. Para anotar semanticamente as regiões, utilizou-se a FMA e a UMLS.

4.4.1.3 Repositório de documentos

Este componente representa os repositórios a partir dos quais o conhecimento é recuperado. No modelo desenvolvido, os repositórios de documentos são caracterizados por possuírem uma camada semântica que permite formalizar a informação contida nos documentos (*i.e.*, textos, imagens, etc.).

Como se pode apreciar na figura 27, os repositórios são definidos em quatro níveis: i) nível descritivo, ii) nível dos metadados,

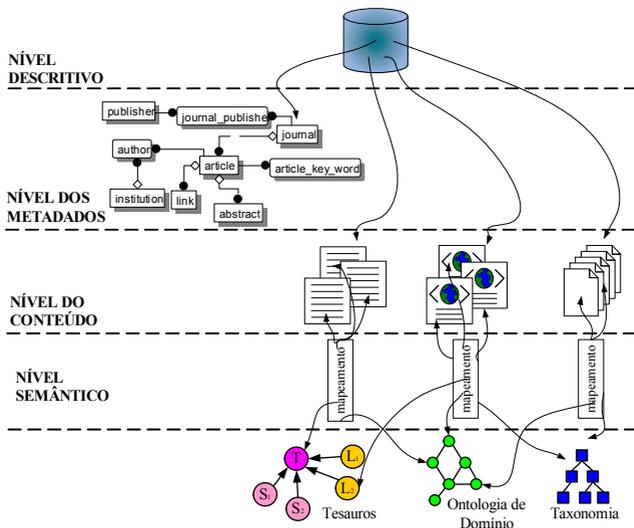


Figura 27 – Estrutura dos repositórios de documentos.

iii) nível do conteúdo e iv) nível semântico. O **nível descritivo** visa identificar o repositório, usando atributos de identificação (*i.e.*, código, nome, descrição, formatos de documento suportados, etc.), atributos físicos (*i.e.*, localização do repositório, etc.) e atributos bibliográficos (*i.e.*, nome, autor/criador, data de criação, versão, etc.). O **nível dos metadados** visa descrever a estrutura da informação armazenada. A estrutura da informação varia de repositório a repositório. Por exemplo, em um repositório de artigos acadêmicos os metadados permitem descrever o título do artigo, os autores, as palavras-chave, etc.; em um repositório de currículos permite descrever a formação acadêmica, a atuação profissional, os projetos desenvolvidos, etc.; em um repositório de prontuários médicos permitiriam descrever o paciente, o médico, sintomas, etc. O **nível do conteúdo** fornece as estruturas necessárias para armazenar os documentos e os metadados associados. Assim como as imagens semânticas, o repositório de documentos também possui um **nível semântico** que visa associar conteúdo semântico aos documentos.

4.4.1.4 Repositório de conhecimento

O repositório de conhecimento é o componente que representa os artefatos usados para representar o conhecimento do domínio usado para i) realizar o mapeamento semântico das imagens e os documentos dos repositórios e ii) realizar as inferências, quando possíveis, sobre os conceitos usados nos processos de visualização do conhecimento. Os tipos de artefatos suportados pelo modelo são vocabulários controlados, glossários, taxonomias, tesouros e ontologias.

4.4.1.5 Visualização

O componente de **Visualização** é o encarregado de apresentar visualmente o conhecimento ao usuário. Este componente baseia-se nas tarefas definida por Shneiderman (1996) para visualizar informação: primeiro obter uma visão global dos dados (*overview first*), concentrar-se em itens de interesse e filtrar itens irrelevantes, (*zoom and filter*) e finalmente fornecer detalhes sob demanda (*then details on demand*).

A visão global dos dados é feita usando as imagens semânticas, os elementos visuais das imagens são alteradas com o intuito de que o usuário facilmente possa identificar as regiões da imagem que se encontram relacionados com o conhecimento recuperados a partir dos repositórios. A filtragem dos itens irrelevantes é facilitada pelos metadados dos repositórios e pelos mapeamentos semânticos feito nos documentos e nas imagens. Para obter os detalhes sob demanda, o componente interage com os repositórios de documentos a fim de obter os itens de informação específicos.

4.4.2 AMBIENTE

A **Necessidade de informação** é considerada como parte do ambiente pois esta não é parte do modelo propriamente dito mas influencia diretamente nos **Usuários**. É esta necessidade que leva aos **Usuários** a procurar documentos nos repositórios digitais para satisfazer seus requerimentos de informação. Estes requerimentos de informação são cada vez mais importantes na sociedade do conhecimento onde o conhecimento é uma das principais fontes de riquezas.

No modelo, a **Necessidade de informação** é o elemento que move os **Usuários** a procurar **Imagens semânticas**. Será por meio

destas **Imagens semânticas** que os requerimentos de informação dos **Usuários** serão satisfeitos. Ao precisar das **Imagens semânticas**, os **Usuários** iniciam um processo onde interagem com o componente de **Visualização** com o objetivo de procurar a imagem que contenha conceitos representados graficamente que encontram-se associados às necessidades de informação. Usando as **Imagens semânticas** os **Usuários** alteram os parâmetros de visualização restringindo assim os resultados das buscas até encontrarem os documentos que precisam.

4.4.3 ESTRUTURAS

4.4.3.1 Visão do mundo dos usuários

A **Visão de mundo** é o elemento que liga os **Usuários** com as **Imagens semânticas**. Esta **Visão de mundo** fornece os modelos mentais para os **Usuários** compreenderem as representações visuais contidas nas imagens. Ao instanciar o modelo, é de suma importância identificar o grupo-alvo, pois somente ao estudar o grupo-alvo será possível identificar a sua **Visão de mundo**. Esta **Visão de mundo** tem que ser considerada no momento da criação das **Imagens semânticas**, a seleção das representações visuais e as anotações semânticas serão feitas tendo em consideração esta visão.

4.4.3.2 Anotações semânticas

Para que o conhecimento seja recuperado de forma eficiente (*i.e.*, buscas semânticas), tanto as imagens quanto os repositórios de documentos devem ter sido previamente enriquecidos com conteúdo semântico obtido a partir do repositório de conhecimento (*i.e.*, ontologias, taxonomias, tesouros). Os mapeamentos semânticos das imagens são realizados de forma manual tendo em consideração a necessidade e visão de mundo do grupo alvo.

Anteriormente, na figura 26 apresentou-se um exemplo de uma imagem usado para visualizar o conhecimento relacionado ao coração onde cada região foi vinculada a uma classe da ontologia FMA incorporando à imagem todo o conhecimento explicitado na ontologia. Por exemplo, ao vincular a região R1 (etiquetada com veia cava superior) à classe **Superior vena cava**, a imagem semântica agora “entende” que aquela região é parte do sistema cardiovascular (classe **Cardiovascular**

system na FMA) e que também é conhecida como *Anterior vena cava* (sinônimo em inglês), *Vena cava superior* (nome equivalente em espanhol) e *Veine cave supérieure* (nome equivalente em francês). Na Figura 26 mostra também como a região R4 (etiquetada como aorta) é mapeada ao conceito C0003483 da UMLS. Usando as relações UMLS, a imagem semântica “sabe” que a *Aneurysm* (código C0002940 na UMLS) é uma doença relacionada à aorta.

BONAMIGO, Telmo Pedro and **SIQUEIRA,** Iara. **Screening for abdominal aortic aneurysms** . Rev. Hosp. Clin. [online]. 2003, v. 58, n. 2, pp. 63-68. ISSN 0041-8781.

OBJECTIVE AND METHODS : Screening for abdominal aortic aneurysms may be useful to decrease mortality related to rupture . We conducted a study to assess the prevalence of abdominal aortic aneurysms in southern Brazil and to define risk factors associated with high prevalence of this disorder . The screening was conducted using abdominal ultrasound . Three groups were studied : Group 1 - cardiology clinic patients ; Group 2 - individuals with severe ischemic disease and previous coronary surgery , or important lesions on cardiac catheterism ; Group 3 - individuals without cardiac disease selected from the general population . All individuals were male and older than 54 years of age . The ultrasonographic diagnosis of aneurysm was based on an anteroposterior **abdominal aorta** diameter of 3 cm, or on an **abdominal aorta** diameter 0.5 cm greater than that of the **supra-renal aorta** .
RESULTS: A total of 2.281 people were screened for abdominal aortic aneurysms in all groups: Group 1 - 768 individuals , Group 2 ...

Ontologia de Referência (FMA)

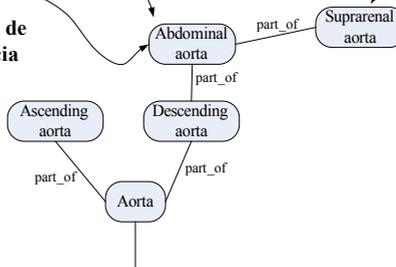


Figura 28 – Exemplo de mapeamento semântico de um documento.

Na figura 28 apresenta-se um exemplo de mapeamento do conteúdo de um documento usando também a ontologia FMA. Os benefícios que o mapeamento semântico oferece aos repositórios de documentos são os mesmo que os fornecidos às imagens semânticas (*i.e.*, formalização do conhecimento, recuperação de informações relacionadas, execução de inferências, etc.). Para o modelo, o fato das imagens e os repositórios de documentos serem mapeados usando os mesmos artefatos de conhecimento permite que estes dois componentes sejam integrados por meio das representações do conhecimento. Por exemplo, na figura 26 a região R4 (etiquetada com aorta) encontra-se mapeada à classe *Aorta* da FMA, por outro lado na Figura 28, observa-se como os conceitos *Abdominal aorta* e *suprarenal aorta* foram identificados no texto de um documento. Apesar que os mapeamentos não foram

feitos usando as mesmas classes FMA, é possível inferir que o documento em questão possui conhecimento relacionado à região R4, já que a aorta supra-renal (classe **suprarenal aorta** na FMA) é parte da aorta abdominal (classe **Abdominal aorta** na FMA) que por sua vez é parte da aorta descendente (classe **Descending aorta** na FMA) que forma parte da Aorta.

4.4.4 MECANISMOS

4.4.4.1 Visualização das imagens

O processo inicia quando o usuário precisa satisfazer alguma necessidade por informação. O usuário seleciona a imagem (número 1 na figura 29) a partir do qual será visualizado o conhecimento. O critério de seleção dependerá diretamente das necessidades de informação do usuário, levando em consideração os conceitos representados nas imagens. Por exemplo, caso seja preciso visualizar o conhecimento associado ao coração, o usuário deverá selecionar uma imagem no qual o coração encontre-se representado graficamente.

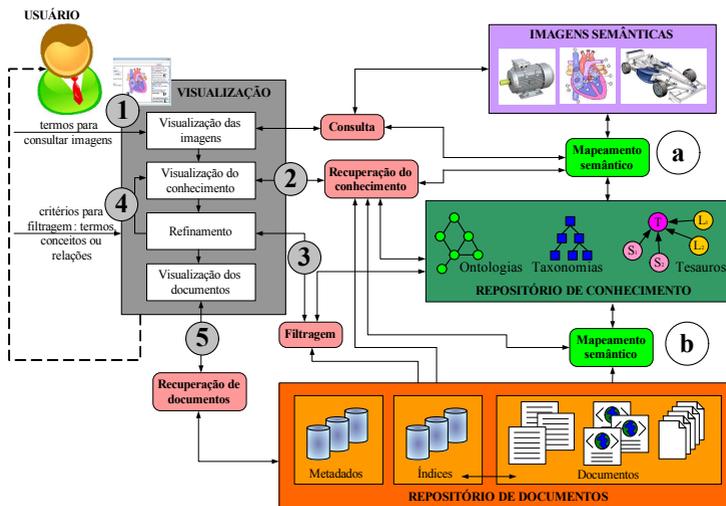


Figura 29 – Modelo proposto para a visualização do conhecimento.

4.4.4.2 Visualização do conhecimento

Uma vez selecionada a imagem, o processo de visualização do conhecimento é executado (número 2 na Figura 29). O processo de visualização visa fornecer ao usuário o conhecimento armazenado nos repositórios de documentos. A visualização dos resultados é feita nas imagens semânticas, de forma que os usuários rapidamente percebam a quantidade de documentos associado a cada região da imagem. Esta representação é refletida na imagem, por exemplo, alterando a coloração das regiões onde se concentra a maior/menor quantidade de documento ou incluir a quantidade de documentos recuperados sobre as regiões.

4.4.4.3 Refinamento

O usuário a partir da visualização fornecida pode restringir o espaço das buscas (número 3 na Figura 29) podendo usar para isto os metadados fornecidos pelos repositórios de documentos ou os conceitos definidos nos artefatos de conhecimento, iniciando assim um novo processo de visualização (número 4 na Figura 29).

4.4.4.4 Visualização dos documentos

Após a execução da **Visualização do conhecimento** os **Usuários** podem recuperar um documento específico (número 5 na Figura 29).

4.5 CONSIDERAÇÕES FINAIS

O objetivo deste capítulo foi apresentar o modelo proposto para a visualização do conhecimento baseado em imagens semânticas. Recorreu-se às áreas de visualização da informação e visualização do conhecimento para compreender as habilidades de temos os humanos para processar as representações visuais e como estas podem auxiliar na divulgação do conhecimento. Para integrar o conhecimento nas imagens semânticas e os repositórios de documentos recorreu-se às anotações semânticas.

O modelo é composto por cinco componentes: os usuários, as imagens semânticas, os repositórios de documentos, o repositórios de

conhecimento, e o componente de visualização. De forma geral o modelo descreve as interações dos usuários com as imagens semânticas a fim de satisfazer uma necessidade de informação. A visão de mundo é usada para criar as imagens semânticas com o objetivo que os usuários compreendam rapidamente a informação visual representados neles e possam relacionar conhecimentos prévios. Para que o modelo compreenda os conceitos representados nas imagens e os mencionados nos documentos foram usados mapeamentos semânticos.

As interações dos usuários são feitas através do componente de visualização. É neste componente que são acionados os mecanismos do modelo: a visualização das imagens permite ao usuário selecionar e visualizar uma imagem, a visualização do conhecimento apresentará no própria imagem os resultados das buscas dos conceitos associados à imagem, o refinamento permite ao usuário se concentrar na informação do seu interesse, e a visualização dos documentos permite que o usuário visualize uma determinada peça da informação.

Nro	Tarefa	Executada por	Conhecimento usado	Intensivo?
1	Ingressar termos para busca	usuário	conhecimento do domínio	não
2	Consultar imagens	agente	conhecimento do domínio	sim
			entidades representadas nas imagens	
3	Obter as entidades representadas na imagem	agente	conhecimento do domínio	sim
			entidades representadas nas imagens	
4	Recuperar os documentos que mencionam as entidades representadas na imagem	agente	conhecimento do domínio	sim
			entidades representadas nas imagens	
			entidades mencionadas nos documentos digitais	
5	Alterar as regiões da imagem onde existem documentos	agente	entidades representadas nas imagens	sim
			entidades mencionadas nos documentos digitais	
6	Apresentar mensagem para o usuário	agente		não

Tabela 18 – Planilha OM-3 CommonKADS para o processo Visua-

Nro	Tarefa	Executada por	Conhecimento usado	Intensivo?
1	Selecionar uma entidade da imagem	usuário	conhecimento do domínio	não
2	Consultar por relações do conceito representado na entidade	agente	conhecimento do domínio	sim
			entidades representadas nas imagens	
3	Apresentar as relações encontradas	agente		não
4	Selecionar relações para filtrar	usuário	conhecimento do domínio	não
5	Recuperar os documentos que mencionam as entidades representadas na imagem e às relações selecionadas	agente	conhecimento do domínio	sim
			entidades representadas nas imagens	
			entidades mencionadas nos documentos digitais	
6	Alterar as regiões da imagem onde existem documentos	agente	entidades representadas nas imagens	sim
			entidades mencionadas nos documentos digitais	

Tabela 19 – Planilha OM-3 CommonKADS para o processo Recupera-

Pergunta chave do <i>framework</i> de Burkhard	Resposta do modelo
Que tipo de conhecimento precisa ser visualizado?	O conhecimento armazenado nos repositórios digitais de documentos.
Por que o conhecimento deve ser visualizado?	Para facilitar a divulgação do conhecimento.
A quem está sendo destinado?	A um grupo de pessoas que compartilham uma mesma visão do mundo.
Em que contexto deveria ser visualizado?	Em um entorno virtual.
Como pode o conhecimento ser visualizado?	Por meio de imagens que contêm representações visuais que <i>a priori</i> são conhecidas pelo grupo-alvo.

Tabela 20 – Aplicação do *framework* para visualização do conhecimento.

Composição	Ambiente	Estrutura	Mecanismos
Usuários	Necessidade de informação	Visão do mundo dos usuários	Visualização das imagens
Imagens semânticas		Anotações semânticas	Visualização do conhecimento
Repositórios de documentos			Refinamento
Repositório de conhecimento			Visualização dos documentos
Visualização			

Tabela 21 – Visão esquemática do modelo.

5 IMPLEMENTAÇÃO DO MODELO

5.1 INTRODUÇÃO

Visando demonstrar a viabilidade de modelo proposto, foi desenvolvido um protótipo aplicado ao domínio biomédico. O repositório de conhecimento foi composto pela ontologia Foundational Model of Anatomy (FMA), uma ontologia de referencia no domínio anatômico (ROSSE; MEJINO, 2003, 2008), e pelo metatesauro Unified Medical Language System (UMLS), um repositório que integra vários vocabulários biomédicos (BODENREIDER, 2004).

Dado que a FMA foi desenvolvida usando frames de Protégé. Usou-se a API do Protégé para acessar as suas classes e propriedades. A UMLS, por outro lado, foi instalada em um banco de dados relacional. Como repositório de documentos usou-se uma cópia da base de dados SciELO. SciELO é uma biblioteca eletrônica que abarca uma coleção seleta de revistas científicas brasileiras. Para anotar os documentos usou-se Metamap, um programa que identifica conceitos UMLS em textos escritos em linguagem natural. Metamap usa uma abordagem intensiva de conhecimento baseada no processamento simbólico, processamento da linguagem natural e técnicas de lingüística computacional (ARONSON, 2001). As imagens semânticas foram implementadas usando a tecnologia Java 2D. Para segmentar as regiões das imagens usou-se AnnotaImage, um software que permite anotar imagens usando a linguagem IML (*Image Markup Language*). AnnotaImage, disponibiliza um XML por cada imagem segmentada, onde as regiões segmentadas são explicitamente especificadas. A anotação das imagens foi feita manualmente vinculando as regiões a conceitos da ontologia FMA e do metatesauro UMLS. Visando facilitar o processo de busca, todos os artigos foram indexados usando Lucene. Tanto o conteúdo textual (título do artigo, resumo e palavras-chaves) quanto as anotações de 77.461 artigos foram indexadas.

Este capítulo esta organizado da seguinte forma após a introdução, se apresenta a arquitetura do protótipo criada para implementar o modelo proposto no capítulo anterior. Posteriormente se descreve cada uma das camadas desta arquitetura. Finalmente são apresentadas as considerações finais deste capítulo.

5.2 ARQUITETURA

O modelo proposto tem sido implementado em uma arquitetura baseada em cinco camadas (ver figura 30): a camada de conhecimento, a camada de anotação, a camada de suporte, a camada de recuperação e a camada de visualização.

A camada de conhecimento implementa o componente repositório do conhecimento do modelo que tem por objetivo acessar às diferentes representações de conhecimento suportadas pelo protótipo. A camada de anotação implementa as principais estruturas do modelo, as anotações semânticas. Esta camada visa fornecer ao protótipo informação semântica sobre o conteúdo das imagens e dos documentos digitais. A camada de suporte implementa os componentes imagens semânticas, repositórios de documentos e os mecanismos, fornece o suporte para as tarefas de visualização e recuperação. A camada de recuperação implementa os mecanismos que permitem a visualização do conhecimento, recuperação de documentos e refinamento das consultas. Finalmente na camada de visualização se implementa o componente de visualização com seu respectivo mecanismo visualização das imagens. Na seqüência, descreve-se a implementação de cada uma destas camadas.

5.3 CAMADA DE CONHECIMENTO

A camada de conhecimento é a camada semântica da arquitetura. Ela contém a ontologia FMA e o vocabulário UMLS. Estas representações de conhecimento são usadas tanto na anotação semântica quanto no processo de recuperação de conhecimento. Cada ontologia é controlada pelo componente **Gestor de Ontologias** o qual usa *wrappers* com o objetivo de integrar estas representações de conhecimento. Desta maneira existirá um *wrappers* por cada representação do conhecimento usada na arquitetura. Os *wrappers* tem sido usadas como meio para integrar múltiplas representações de conhecimento encapsulando os origens e provendo mecanismos que são independentes da estrutura da fonte (STUDER et al., 2000; BRINKLEY et al., 2006).

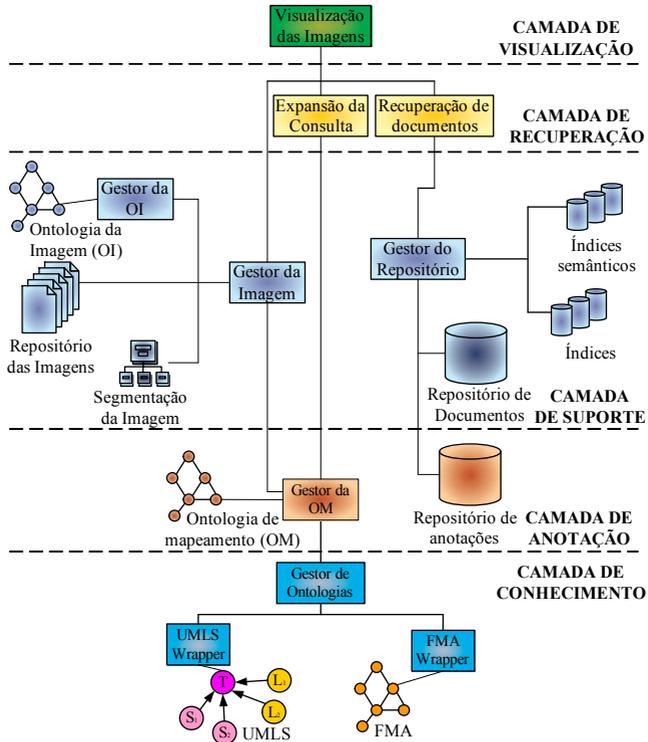


Figura 30 – Arquitetura do protótipo.

5.3.1 FMA Wrapper

Um modelo de referência pode ser definido como a conceitualização do corpo de conhecimento relativo a um domínio, no qual se declara as restrições requeridas para incluir os conceitos e as relações, explicitamente se definem os conceitos e também cria, de forma explícita, as relações entre os conceitos que são tanto necessários e suficientes para a modelagem da estrutura de forma coerente e consistente do domínio do conhecimento específico (ROSSE; SHAPIRO; BRINKLEY, 1998). Neste sentido a FMA (*Foundational Model of Anatomy*) é um modelo de referência sobre a anatomia e no qual se encontra representado de forma simbólica a estrutura organizacional do corpo humano desde o nível macromolecular até o nível macroscópico. Esta tem sido concebida para ser usada em diferentes problemas que requerem conhecimento anatômico

(ROSSE et al., 1998; ROSSE; MEJINO, 2003). A FMA tem sido implementada usando frames de Protégé e é armazenada em um banco de dados relacional MySQL. Contem aproximadamente 75,000 classes e aproximadamente 2.1 milhões de instâncias de relações.

O FMA *Wrapper* tem sido implementado usando o API do Protégé¹. A base de conhecimento é instanciada a partir do arquivo de projeto da FMA usando as classes `Project` e `KnowledgeBase`. Dado que a FMA está estruturada em *frames*, os métodos para recuperar as classes e propriedades são baseados principalmente em dois métodos `getFrame` e `getSlot`. O primeiro destes é usado para recuperar as classes e o segundo para recuperar as suas propriedades. Na seqüência, apresenta-se parte do código usado para recuperar a base de conhecimento FMA, uma classe FMA e as propriedades de uma classe.

Listagem 5.1 – Implementação do *wrapper* para a FMA

```
private KnowledgeBase oProtege;
private Project oProjet;
...
public Frame getFMAKnowledgeBase(String FileName)
{
    Collection<Object> errors;
    errors = new ArrayList<Object>();
    this.oProject = new Project(FileName, errors);
    this.oProtege = this.oProject.getKnowledgeBase();
}

public Frame getFMAClass(String nmClass)
{
    Frame frame = this.oProtege.getFrame(nmClass);
    return frame;
}

public Iterator<Frame> getFMAProperties(Frame frame,
                                       String nmProperty)
{
    Slot slot = oProtege.getSlot(nmProperty);
    Collection<Frame> col;
    col = frame.getOwnSlotValues(slot);
    return col.iterator();
}
```

¹<http://protege.stanford.edu/>

5.3.2 UMLS Wrapper

A UMLS é um repositório de vocabulários biomédicos que integra aproximadamente 5 milhões de nomes para 1 milhão de conceitos de mais de 100 famílias de vocabulários biomédicos, assim como 16 milhões de relações entre estes conceitos (BODENREIDER; ZHANG, 2006). A UMLS é composta por três componentes: o meta-tesauro, a rede semântica e os recursos léxicos (BODENREIDER, 2004). O meta-tesauro é o principal componente da UMLS, este contém informação acerca de conceitos relacionadas à biomedicina e à saúde, seus sinônimos, e as relações entre eles. A rede semântica fornece um alto nível de classes para categorizar os conceitos do meta-tesauro. Os dados da UMLS são armazenados em um conjunto de tabelas relacionais e arquivos.

Na implementação do modelo tem sido usada a UMLS armazenada em um banco de dados relacional. O UMLS Wrapper foi implementado usando consultas SQL. O banco de dados UMLS possui uma conjunto de tabelas que armazenam os conceitos dos diversos vocabulários, as relações entre estes conceitos, os tipos semânticos das relações, entre outras informações.

A UMLS é baseada em conceitos, assinando um código único que identifica ao conceito dentro do banco de dados. Um conceito pode estar definido em várias ontologias e vocabulários, mas compartilham o mesmo código. Por exemplo, os conceitos **heart** (coração em inglês), **corazón** (coração em espanhol), e **coração** têm o mesmo código no banco de dados UMLS: C0018787. Na seqüência, apresenta-se um exemplo de uma consulta SQL executada pelo UMLS Wrapper com o objetivo de recuperar todos os sinônimos do conceitos coração.

Listagem 5.2 – Implementação do *wrapper* para a UMLS

```
select eCON.cui, eCON.str, eCON.lat,
       eCON.lui, eCON.sui, eCON.aui
from MRCONSO eCON
where CUI = "C0018787"
```

5.4 CAMADA DE ANOTAÇÃO

Existem dos elementos que necessitam ser anotados na arquitetura: o repositório das imagens e o repositório dos documentos. Para que a arquitetura possa recuperar os documentos a partir das imagens, esta deveria “entender” o conteúdo representado tanto nas imagens

quanto nos documentos. Por isto é necessário enriquecer estas estruturas para incluir conteúdo semântico. Isto é feito através de processos de anotação.

No protótipo, a estrutura que suporta às anotações semântica das imagens é a **Ontologia de Mapeamento**, a qual é acessada pelo **Gestor da OM**. Por outro lado, as anotações dos documentos digitais são armazenadas na arquitetura no **Repositório de anotações**. Na seqüência, descreve-se os processos de anotação semântica tanto das imagens quanto dos documentos.

5.4.1 Anotações Semânticas das Imagens

O mapeamento de ontologias pode ser entendida como o problema de encontrar o mapeamento semântico entre duas ou mais ontologias (DOAN et al., 2002, 2004; EHRIG; SURE, 2004; NOY; DOAN; HALEVY, 2005; CHOI; SONG; HAN, 2006; SHVAIKO; EUZENAT, 2008). De formar geral o mapeamento pode ser especificado como uma função que transforma instâncias de uma ontologia em instâncias da outra (CALVANESE; LENZERINI, 2001). Nas abordagens habituais para mapear ontologias, os mecanismos para especificar o mapeamento entre conceitos de ontologias diferentes são limitados para expressar correspondência direta entre os termos (CALVANESE; LENZERINI, 2001). Infelizmente não existem padrões para mapear ontologias (EHRIG, 2005), mas ontologias de mapeamento tem sido usadas como formas de representação (MAEDCHE et al., 2002; CRUBÉZY; MUSEN, 2003; NOY et al., 2004).

Na arquitetura se usa o mapeamento entre as ontologias para integrar nas imagens os conceitos definidos nas representações do conhecimento da camada de conhecimento. As instâncias desta ontologia representarão o mapeamento entre as entidades da imagem com as instâncias, classes ou conceitos das representações do conhecimento (ver figura 31). Assim por exemplo, uma instância desta ontologia poderia indicar que a entidade **Aorta** de uma imagem encontra-se mapeada à classe **Aorta** da FMA e ao termo **Aorta** UMLS.

Na figura 32 pode se apreciar as principais classes da **Ontologia de mapeamento**. As diversas representações de conhecimento são definidas como instâncias da classe **Information Source**, assim por exemplo esta classe poderia ter como instâncias à ontologia FMA, ao vocabulário UMLS, à Disease Ontology, entre outros. Os elementos das representações do conhecimento (*i.e.*, URI da classe ou instancia, código do termo, palavra-chave ou conceito) são representados como instâncias

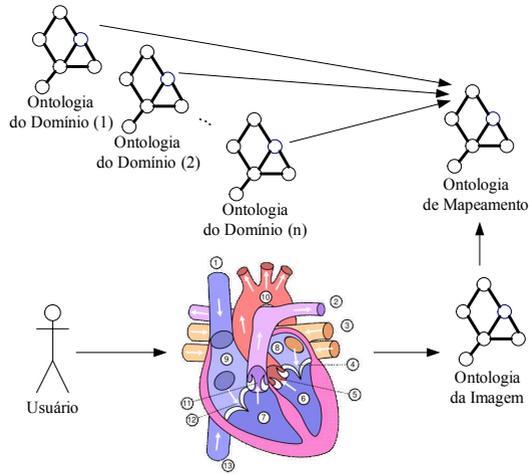


Figura 31 – As imagens semânticas e sua integração com o conhecimento do domínio.

da classe `InformationElement`.

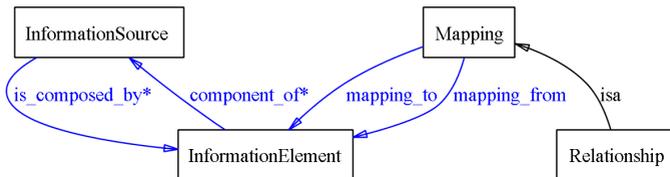


Figura 32 – Ontologia de mapeamento.

O mapeamento é definido na classe `Mapping` que se baseia nas relações `mapping-from` e `mapping-to`. A relação `mapping-from` usa como elementos de informação instâncias da classe `Concept` da *Ontologia da imagem*. Os tipos de informação usados pela relação `mapping-to` dependerão do tipo de fonte podendo ser estas classes, instâncias e/ou conceitos. A classe `Relationship` permite identificar a relação que existem entre os elementos de informação.

A figura 33 mostra um exemplo de anotação da região de uma imagem, a região rotulada como `aorta` tem sido associada ao conceitos C0003483 da UMLS e 3734 da FMA. Adicionalmente, também se apresenta uma instância de mapeamento diferente: *related diseases*

(mapping-rel). Esta ontologia permite definir outros tipos de relações de mapeamentos, fornecendo uma grande flexibilidade na hora de anotar as regiões das imagens.

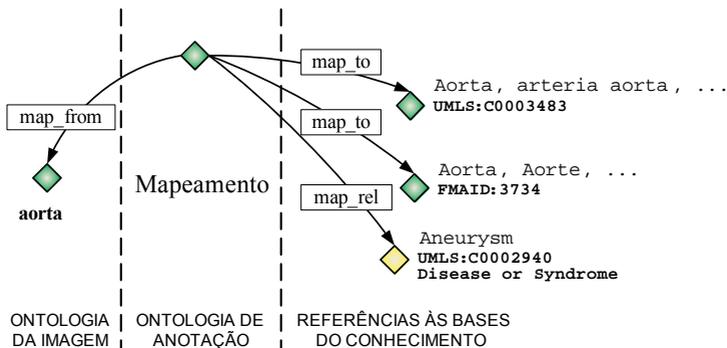


Figura 33 – Esquema de anotações das imagens.

5.4.2 Anotações Semânticas dos Documentos

Na arquitetura proposta as entidades mencionadas nos documentos digitais têm sido armazenadas no **Repositório de anotações** visando depois a criação dos **Índices semânticos**. Na figura 34 apresenta um exemplo de como a arquitetura implementa este processo. Na figura 34-(a) apresenta parte de um documento sendo anotado, as entidades são reconhecidas no conteúdo do documento e é criada a representação do documento anotado semanticamente, como observa-se na figura 34-(b). Usando esta representação a arquitetura gera os índices semânticos para cada documento (figura 34-(c)). A arquitetura não se encarrega do processo de anotação, as tarefas de reconhecimento de entidades podem ser feitas usando técnicas da área extração de informação (CUNNINGHAM, 1997; MAYNARD; BONTCHEVA; CUNNINGHAM, 2003; CUNNINGHAM, 2005).

Cada artigo armazenado na base de dados SciELO tem sido pré-processado com o objetivo de incluir anotações para identificar os conceitos dentro dos textos. Para identificar os conceitos UMLS foi usado MetaMap. Metamap é um programa cujo objetivo é identificar conceitos UMLS dentro de textos biomédicos para o qual usa uma abordagem baseada no processamento simbólico, processamento da linguagem natural e técnicas da lingüística computacional (ARONSON, 2001).

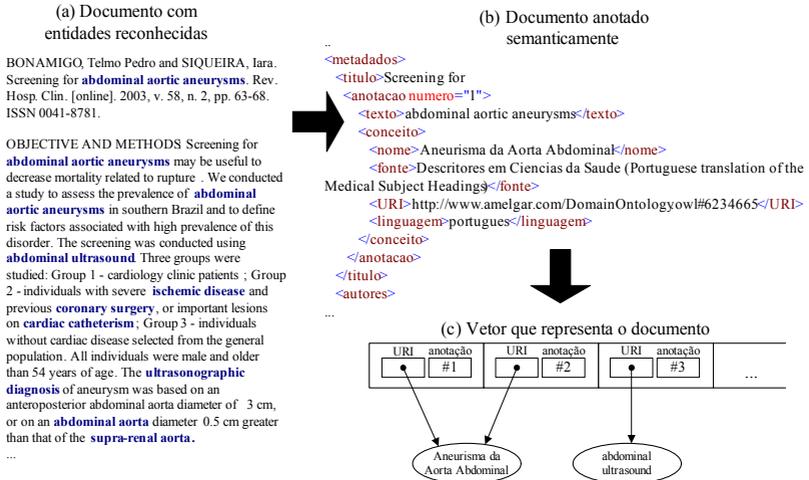


Figura 34 – Anotações dos documentos digitais.

Na versão implementado do protótipo, somente as títulos, resumos e palavras-chaves foram anotados. Na seqüência, apresenta-se parte de uma titulo anotado semanticamente usando o MetaMap.

Listagem 5.3 – Resumo do Artigo “*Screening for abdominal aortic aneurysms*” (BONAMIGO; SIQUEIRA, 2003) anotado usando o formato XML do MetaMap

```

<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE MMOs PUBLIC "-//NLM//DTD MetaMap Machine Output
//EN" "http://ii-public.nlm.nih.gov/DTD/MMOtoXML_v4.
dtd">
<MMOs>
<MMO>
...
<SyntaxUnit>
<SyntaxType>head</SyntaxType>
<LexMatch>abdominal aortic aneurysms</LexMatch>
<InputMatch>abdominal aortic aneurysms</InputMatch>
<LexCat>noun</LexCat>
<Tokens Count="3">
<Token>abdominal</Token>
<Token>aortic</Token>

```

```

    <Token>aneurysms</Token>
  </Tokens>
</SyntaxUnit>
</SyntaxUnits>
<PhraseStartPos>10</PhraseStartPos>
<PhraseLength>30</PhraseLength>
<Candidates Count="12">
  <Candidate>
    <CandidateScore>-1000</CandidateScore>
    <CandidateCUI>C0162871</CandidateCUI>
    <CandidateMatched>Abdominal Aortic Aneurysms</
      CandidateMatched>
    <CandidatePreferred>Aortic Aneurysm, Abdominal</
      CandidatePreferred>
    <MatchedWords Count="3">
      <MatchedWord>abdominal</MatchedWord>
      <MatchedWord>aortic</MatchedWord>
      <MatchedWord>aneurysms</MatchedWord>
    </MatchedWords>
    <SemTypes Count="1">
      <SemType>dsyn</SemType>
    </SemTypes>
    <MatchMaps Count="1">
      <MatchMap>
        <TextMatchStart>1</TextMatchStart>
        <TextMatchEnd>3</TextMatchEnd>
      </MatchMap>
    </MatchMaps>
  </Candidate>
  ...
</Utterances>
</MMO>
</MMOs>

```

Os resultados oferecidos por Metamap contêm uma ampla variedade de elementos que podem ser usados em diversas aplicações da engenharia do conhecimento. Em particular, para a implementação do protótipo, interessa basicamente a identificação das entidades UMLS. Por este motivo o XML retornado pelo Metamap foi adaptado para conter somente os elementos de interesse para o protótipo. Além disso, usar uma representação própria das anotações dos documentos permitiria integrar mapeamentos de outras ferramentas em futuras versões do protótipo. Na seqüência, apresenta-se parte de uma título anotado semanticamente adaptado do Metamap.

Listagem 5.4 – Resumo do Artigo “*Screening for abdominal aortic aneurysms*” (BONAMIGO; SIQUEIRA, 2003) anotado adaptando o formato XML do MetaMap

```

<?xml version="1.0" encoding="ISO-8859-1"?>
<Document>
<Phrase>
...
<PhraseText>for abdominal aortic aneurysms</PhraseText>
<Annotations><Annotation Type="UMLS">
<Mapping CandidateMatched="Abdominal Aortic Aneurysms"
CandidatePreferred="Aortic Aneurysm, Abdominal" score=
"-1000">C0162871</Mapping>
<SemTypes>
<SemType>dsyn</SemType>
</SemTypes>
</Annotation>
<Annotation Type="UMLS_CANDIDATE"><Mapping
CandidateMatched="Abdominal Aortic Aneurysms"
CandidatePreferred="Aortic Aneurysm, Abdominal" score=
"-1000">C0162871</Mapping>
<SemTypes>
<SemType>dsyn</SemType>
</SemTypes>
</Annotation>
<Annotation Type="UMLS_CANDIDATE">
<Mapping CandidateMatched="Aortic aneurysms"
CandidatePreferred="Aortic Aneurysm" score="-901">
C0003486</Mapping>
<SemTypes>
<SemType>anab</SemType>
<SemType>dsyn</SemType>
</SemTypes>
</Annotation>
<Annotation Type="UMLS_CANDIDATE"><Mapping
CandidateMatched="Abdominal aneurysm"
CandidatePreferred="Aortic Aneurysm, Abdominal" score=
"-896">C0162871</Mapping>
<SemTypes>
<SemType>dsyn</SemType>
</SemTypes>
...

```

</Phrase></Document>

5.5 CAMADA DE SUPORTE

A camada de suporte é a parte da arquitetura que põe a disposição os componentes que serão usados nas tarefas de visualização do conhecimento e recuperação do conhecimento. Esta camada é composta por dois grupos de componentes, uns dedicados à gestão das imagens e outro dedicado à gestão dos repositórios de documentos digitais. Em ambos os casos, a camada de anotação é usada para prover a informação semântica necessária para determinar quais as entidades anatômicas que têm sido representadas em cada imagem e quais os conceitos referenciados nos repositórios digitais.

5.5.1 Gestão das Imagens

A gestão das imagens é controlada pelo componente **Gestor da Imagem**. Este componente usa a **Ontologia da Imagem** na qual se descrevem as imagens. A **Ontologia da Imagem** permite descrever semanticamente quais entidades têm sido representadas em cada região da imagem, identificando a região com seu respectivo nome. Este nome será posteriormente vinculado a um ou mais conceitos das representações de conhecimento.

As imagens que serão instanciadas nesta ontologia são recuperadas a partir de um banco de dados representados na arquitetura pelo componente **Repositório de Imagens**. Antes de incorporar as imagens na ontologia, estas são pré-processadas visando identificar e etiquetar as suas regiões. Este processo é executado pelo componente **Segmentação da Imagem**.

Para segmentar a imagem foi usada a linguagem *Image Markup Language* (IML). A IML é um esquema que permite incluir informação textual nas imagens (LOBER et al., 2001). Tem sido usado em *software* educativos para a medicina e em sistemas de informação clínicos. No modelo de anotação da IML as regiões da imagem são identificadas por etiquetas escritas em linguagem natural. Na seqüência, apresenta-se parte de um arquivo IML usado para segmentar uma imagem que representa um coração.

```

<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<Frame image="coracao.png">
  <Colors outline_color="YELLOW" pin_color="YELLOW"
    string_color="YELLOW"/>
  <Region id="21" type="STRUCT">
    <label>veia pulmonar inferior direita</label>
    <center>
      <pt x="107" y="269"/>
    </center>
    <outline>
      <pt x="143" y="248"/>
      <pt x="140" y="267"/>
      <pt x="136" y="284"/>
      <pt x="132" y="293"/>
      <pt x="81" y="292"/>
      <pt x="73" y="279"/>
      <pt x="72" y="263"/>
      <pt x="75" y="253"/>
      <pt x="82" y="246"/>
    </outline>
  </Region>
  <Region id="20" type="STRUCT">
    <label>veia pulmonar superior direita</label>
    <center>
      <pt x="115" y="213"/>
    </center>
    <outline>
      <pt x="145" y="192"/>
      <pt x="146" y="217"/>
      <pt x="145" y="237"/>
      <pt x="97" y="236"/>
      <pt x="88" y="227"/>
      <pt x="86" y="215"/>
      <pt x="85" y="206"/>
      <pt x="88" y="199"/>
      <pt x="93" y="193"/>
      <pt x="97" y="190"/>
    </outline>
  </Region>
  ...
</Frame>

```

Na figura 35 se pode apreciar as principais classe da **Ontologia da Imagem**. As imagens são representadas nesta ontologia na classe **Image**, O nome físico da imagem, a data de criação e demais metadados de conteúdo independente são representados nesta classe. As entidades existentes em cada imagem são representadas por regiões. Uma região é definida mediante um conjunto de *píxeles* considerando que um *pixel* pode pertencer unicamente a uma região. As regiões são representadas na ontologia como instâncias da classe **ImageRegion**.

Os conceitos, definidos como instâncias da classe **Concept**, são descritos usando linguagem natural. Esta ontologia estes conceitos não se relacionam com os conceitos definidos na camada de conhecimento. A idéia é manter a descrição da imagem e a anotação da imagem em estruturas independentes, isto permite que uma mesma imagem possa ser mapeada com diversas anotações seguindo diferentes critérios.

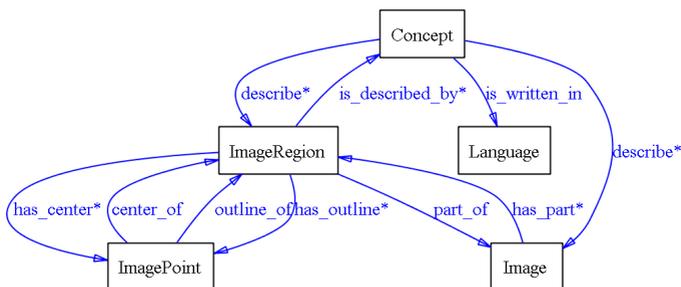


Figura 35 – Ontologia da Imagem.

5.5.2 Repositório dos Documentos Digitais

A gestão dos repositórios é controlado pelo componente **Gestor do Repositório**. Este componente, tendo como entrada o **Repositório de Anotações**, é responsável da criação dos índices que serão usados como as estrutura de suporte no processo de recuperação de documentos. Existem dois tipos de índices, um usado para recuperação de informação textual e outro usado para a recuperação de informação semântica. Os índices são representados na arquitetura pelos componentes **Índices** e **Índices Semânticos**. O componente **Gestor do Repositório** também é o responsável pela recuperação dos documentos a partir do **Repositório de Documentos**.

A arquitetura tem sido implementada usando como repositório de documentos uma cópia da base de dados Scielo. SciELO² (*Scientific Electronic Library Online*) é uma biblioteca eletrônica que abarca uma seleta coleção de periódicos científicos brasileiros. Esta base de dados tem sido modelada para armazenar informação relacionada a artigos acadêmicos (*i.e.*, *title*, *journal*, *publisher*, *authors*, *abstract*, *key word*, etc). O modelo desta base de dados pode ser visualizado na figura 36.

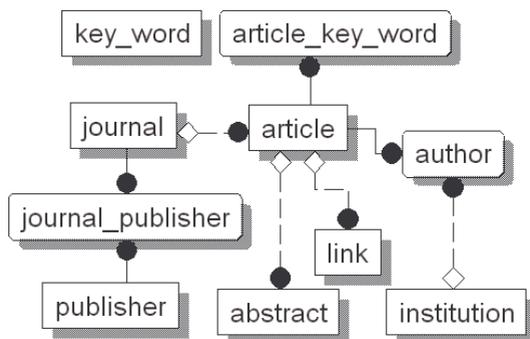


Figura 36 – Organização da base de dados SciELO.

5.6 CAMADA DE RECUPERAÇÃO

A camada de recuperação é responsável pela execução dos processos de recuperação de informação solicitados pela capa de visualização. Dois componentes formam parte desta camada: o componente **Expansão da Consulta** e o componente **Recuperação de Documentos**.

O componente **Expansão da Consulta** é o responsável por enriquecer a consulta, acrescentando termos e/ou conceitos. Para expandir a consulta, este componente se baseia nos conceitos mapeados nas imagens. O componente **Recuperação de Documentos** é responsável pela implementação da recuperação de documentos. Basicamente é uma interface que encapsula as estruturas de recuperação de informação.

²<http://www.scielo.br/>

5.7 CAMADA DE VISUALIZAÇÃO

Esta camada compõe o componente **Visualização das Imagens**. É o responsável de apresentar as imagens aos usuários. Devido a que as imagens têm sido anotadas previamente, este componente pode “entender” o que está representado na imagem. Esta importante característica permite a recuperação de documentos sem que o usuário especifique os termos para a consulta. somente é preciso especificar os termos para procurar as imagens.

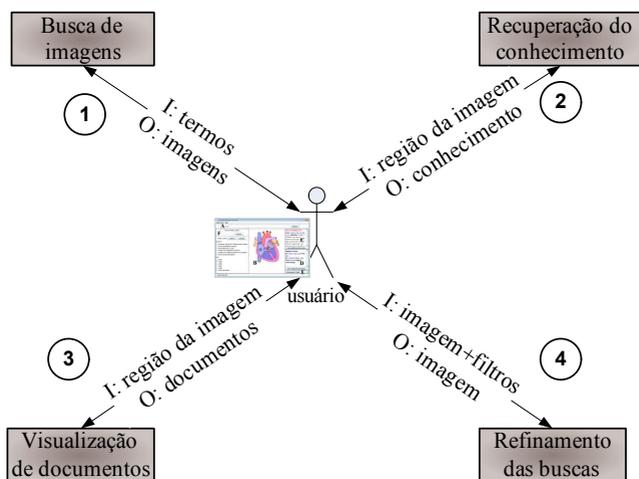


Figura 37 – Processos implementados no protótipo.

As principais operações que o usuário pode executar podem ser visualizadas na figura 37. O processo de recuperação do conhecimento começa com a busca das imagens (figura 37-1). As imagens são usadas como veículos para a recuperação do conhecimento. Quando é selecionada uma região da imagem é recuperado todo o conhecimento relacionado a dita região (figura 37-2). Outra operação que o usuário pode executar sob a região da imagem, é a recuperação de documentos (figura 37-3). Com o objetivo de limitar os resultados das buscas, o usuário pode aplicar tantos os filtros textuais quanto os filtros semânticos (figura 37-4). Esta operação ocasionará que a informação relacionada à imagem seja atualizada.

5.7.1 Busca de imagens

Para procurar as imagens o usuário ingressa os termos de busca na caixa de texto no painel superior (figura 38-A). A ferramenta recupera as imagens associados aos termos processando a consulta tanto no nível estrutural (*i.e.*, busca textual pelo nome das regiões) quanto no nível semântico (*i.e.*, busca conceitual pelo mapeamento). Quando o processamento é feito no nível semântico, os termos são transformados em conceitos do domínio. Esta transformação permite que a consulta seja executada semanticamente, tornando a busca independente da linguagem. Por exemplo, para pesquisar imagens relacionadas ao coração, o usuário pode inserir como termos de pesquisa palavras como coração, *corazón* (coração em espanhol), *heart* ou (coração em inglês). Em todos os casos, o processo retorna os mesmos resultados.

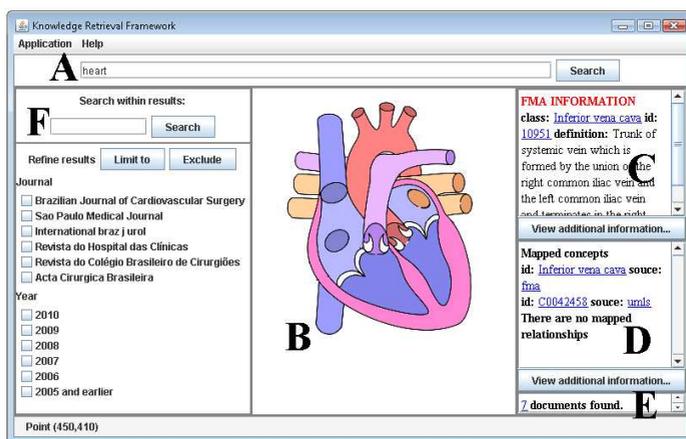


Figura 38 – Protótipo para visualização do conhecimento.

5.7.2 Recuperação do conhecimento

Depois de recuperar a imagem, este é colocado no painel central (figura 38-B). O usuário pode utilizá-lo de duas maneiras: para obter conhecimentos relacionados aos conceitos representados nas regiões ou para recuperar os documentos que mencionam os conceitos.

O conhecimento relacionado ao conceito é apresentado no painel direito no topo (figura 38-C). Nesta versão do protótipo, esta in-

formação é apresentada usando a FMA. Neste painel o usuário pode visualizar o nome, a identificação e a descrição da classe. Usando o botão “*View additional information*”, pode obter informações adicionais como nomes em outras línguas ou as entidades anatômicas que a constituem. Os conceitos mapeados em cada região podem ser observados no painel direito no centro (figura 38-D). Esta informação é obtida a partir do nível semântico da imagem.

5.7.3 Visualização de documentos

Ao selecionar uma região da imagem, além de apresentar ao usuário informações sobre o conceito, também é apresentada a quantidade de documentos relacionados com os conceitos mapeados em cada região. Esta informação se encontra no painel direito, ao selecionar o vínculo (figura 38-E), é apresentado um formulário contendo uma lista de todos os documentos do repositório relacionados a esse conceito.

Na figura 39 pode-se visualizar uma lista dos documentos recuperados. Como cada documento está anotado, vínculos de cada conceito reconhecido nos documentos estão disponíveis, ao selecioná-los uma descrição do conceito é apresentada.

5.7.4 Refinamento das buscas

O usuário pode refinar os resultados da pesquisa utilizando as opções disponíveis no painel esquerdo (figura 38-F). Este painel apresenta duas opções, filtragem por revistas (*i.e.*, journals) e filtragem por anos. O usuário pode procurar por texto dentro de documentos, ou usar os filtros para refinar a busca.

5.8 CONSIDERAÇÕES FINAIS

O objetivo deste trabalho consiste na proposta de um modelo destinado a visualização do conhecimento baseado em imagens semânticas a fim de facilitar a organização ontológica do conhecimento. Com o objetivo de testar o modelo proposto, foi desenvolvido um protótipo usando imagens na área da anatomia e repositórios de documentos no domínio biomédico.

No protótipo proposto tanto as imagens quanto os repositórios de documentos têm sido enriquecidos com conteúdo semântico possibi-

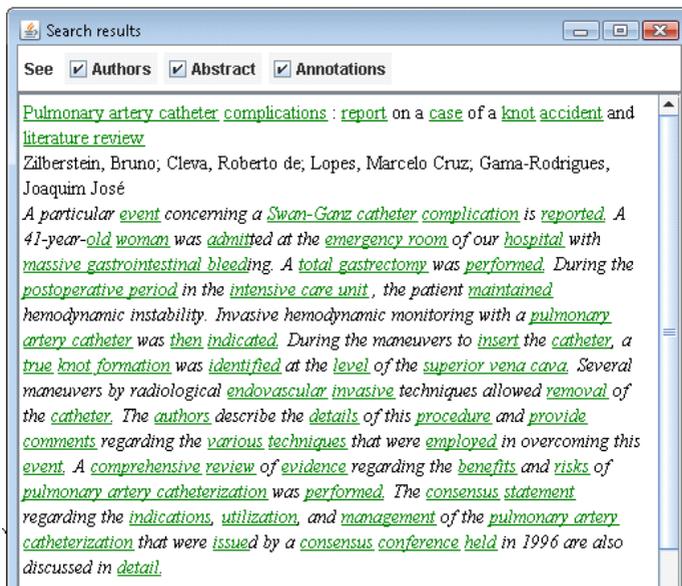


Figura 39 – Visualização dos documentos. No exemplo se visualiza o artigo recuperado de Lopes et al. (2004).

litando a integração destes dois componentes. O conteúdo semântico permite ao protótipo “entender” quais os conceitos representados tanto nas imagens quanto nos documentos facilitando assim a busca conceitual. Uma das vantagens disto é a recuperação de conhecimento de forma independente à linguagem da escrita dos documentos. Devido à este “entendimento”, o protótipo é capaz de recuperar os documentos relacionados aos conceitos representados nas imagens auxiliando o usuário no processo de busca. Este mecanismo de recuperação pode ser observado no protótipo, para recuperar documentos o usuário só precisa selecionar uma região da imagem sem necessidade de ingressar os termos das buscas como nos sistemas de recuperação tradicionais. Entretanto, o usuário deve ingressar termos para procurar imagens. Uma vez selecionado a imagem o modelo “entende” quais as informações necessita recuperar usando para tal os mapeamentos semânticos.

A idéia do modelo é que os resultados das buscas sejam apresentados de forma global na própria imagem alterando a coloração das regiões onde se concentra a maior quantidade de documentos. Este comportamento do modelo é consistente com as tarefas de visualiação

definidas por Shneiderman (1996) (*i.e.*, primeiro apresentar a visão global dos dados, depois concentrar-se em itens de interesse e filtrar itens irrelevantes e finalmente fornecer a informação sob demanda).

O motivo para usar imagens contendo representações do mundo real se baseou por um lado nas habilidades que dos seres humanos em processar imagens rapidamente e por outro lado na facilidade dos indivíduos para relacionar conhecimentos prévios associados a uma imagem já conhecida. Ao ver uma imagem que pertence a sua visão do mundo, as pessoas sabem o que ela representa apesar de não se lembrar dos nomes dos conceitos nela representados.

O protótipo proposto é genérico e pode ser usado em qualquer domínio que permita a representação de conceitos em imagens. Pode ser usado também sobre qualquer repositório de documentos desde que estes possam ser mapeados em representações de conhecimentos, como, por exemplo, ontologias ou taxonomias.

Um dos requerimentos do modelo é a necessidade de mapeamento semântico das imagens e dos repositórios de documentos. Os repositórios de documentos caracterizam-se por armazenar grandes quantidades de informação fazendo com que o processo de anotação manual seja inviável. Por outro lado, a anotação automática ou semi-automática pode ser facilitada pelos métodos da área da extração de informação, mas a precisão destes métodos pode variar dependendo do domínio e tipo de documento. No caso das imagens, as anotações manuais são viáveis na maioria dos domínios já que a quantidade de conceitos que se espera que sejam representados neles é relativamente pequena em comparação com os conceitos representados nos documentos.

Uma das vantagens que oferece o modelo é a facilidade para recuperação de documentos usando apenas regiões das imagens. Nos sistemas de recuperação tradicionais, os usuários traduzem suas necessidades de informação em termos de buscas, estes retornam uma lista de itens que correspondem aos documentos mais relevantes segundo os termos informados. No modelo proposto os usuários traduzem suas necessidades de informação também em termos, mas estes termos não são usados para procurar diretamente os documentos e sim para procurar as imagens que serão usadas no processo de visualização.

Outra vantagem do modelo é a integração dos componentes por meio dos mapeamentos semânticos. As imagens e os diferentes repositórios de informação encontram-se todos integrados por meio de informação semântica, isto permite que uma mesma imagem possa ser usada para visualizar documentos de diferentes repositórios. Por exem-

plô, a imagem do coração apresentado nas seções anteriores, pode ser usado para recuperar artigos científicos, competências organizacionais, projetos, estudos clínicos, indicadores, imagens médicas, entre outros. Esta integração semântica também permite que o modelo possa inferir novas informações sobre um determinado conceito da imagem. Por exemplo, ao projetar o coração, é possível recuperar informações relacionadas a conceitos que não são definidos de forma explícita na imagem, mas que podem ser inferidos, como por exemplo, certas doenças. No coração ao selecionar a aorta, por exemplo, e usando as relações UMLS poderiam ser recuperados os documentos relacionados ao aneurisma, uma doença que afeta a aorta.

6 DEMONSTRAÇÃO DE VIABILIDADE E ANÁLISE COMPARATIVA

6.1 APRESENTAÇÃO DO PROTÓTIPO

A tela inicial do sistema é uma tela bem simples, possui apenas um campo para ingressar os termos para fazer as buscas das imagens e um botão que executa a busca (ver figura 40). Ela tem sido desenhada considerando as recomendações da teoria da carga cognitiva que busca reduzir a carga extrínseca para que o usuário tenha mais capacidade cognitiva livre para construir os seus modelos mentais.

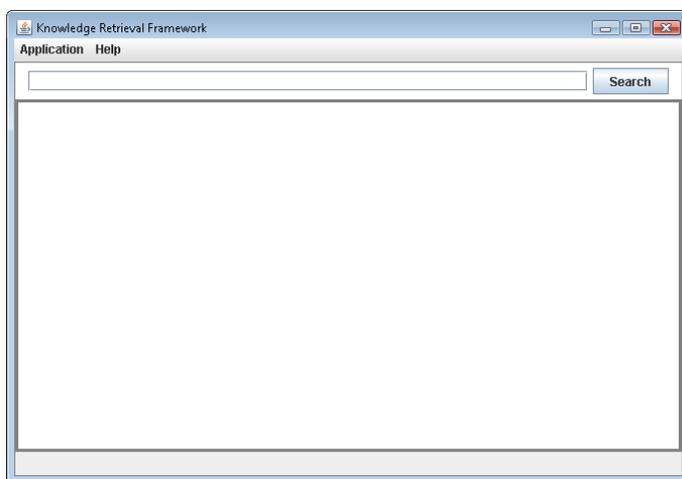


Figura 40 – Tela inicial do protótipo.

Após executada a busca, o painel embaixo da caixa de texto, apresentará a imagem a partir da qual se visualizará o conhecimento recuperado a partir dos repositórios de informação digitais. O conhecimento que vai ser recuperado encontra-se especificado nos mapeamentos feitos nas regiões das imagens.

Para buscar uma imagem, como por exemplo, uma relacionada ao coração, o usuário poderia ingressar como termos para a busca a palavra *coração* ou a palavra *corazón* (coração em espanhol) ou a palavra *heart* (coração em inglês). Isto é possível devido a que o protótipo tem implementado buscas semânticas. Os termos ingressados para as buscas

são previamente traduzidos para encontrar o conceito correspondente nas representações de conhecimento armazenadas na camada de conhecimento. Por exemplo, ao ingressar o termo coração, o protótipo antes de fazer a busca tenta encontrar o conceito coração, na implementação do protótipo isto corresponde ao conceito UMLS C0018787, o mesmo conceito teria sido encontrado com os termos *corazón* e *heart*. É com este conceito que o protótipo faz a busca e é este o motivo pelo qual não interessa a linguagem em que está escrito o termo desde que esteja explicitado nas representações de conhecimento usadas. Na figura 41 apresenta-se a tela que apresenta os resultados da busca pelo termo coração.

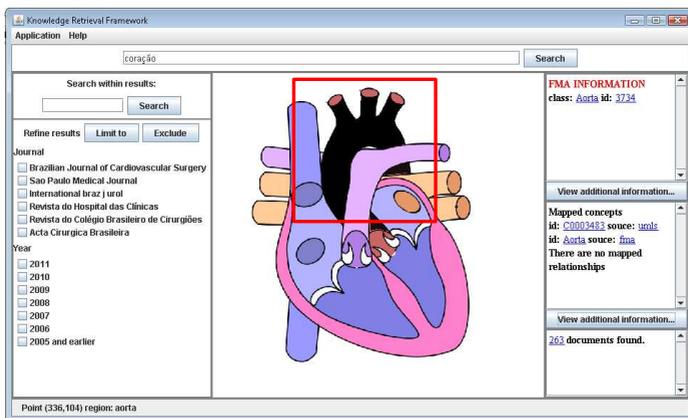


Figura 41 – Tela resultante da busca da imagem pelo termo “coração”.

A visualização do conhecimento é feita usando a mesma imagem recuperada. Na figura 41 apresenta uma visualização que responde a pergunta: *Qual é a entidade mais referenciada nos artigos?* Nesta visualização a região da imagem que representa a entidade com a maior quantidade de documentos é apresentada com uma cor diferente, neste caso o preto. A idéia por trás desta visualização é que o usuário rapidamente identifique esta região. Usou-se a cor por ser um elemento pré-atentivo o qual é identificado rapidamente pelo nosso cérebro. No protótipo implementado foram mapeados artigos científicos, mas poderiam mapear-se outro tipo de informação como, por exemplo, as competências pessoais assim usando esta representação se poderia visualizar a entidade onde a organização tem maior competência ou a entidade onde se tem menor competência.

Usando a imagem recuperada, o usuário pode também visualizar

a definição de cada região. No protótipo implementado tem sido usada a ontologia FMA para este propósito dado que esta ontologia é uma ontologia de referência na área da anatomia. Na figura 42 pode se observar como ao selecionar a região *veia cava inferior* no painel da direita aparece a definição de tal entidade. As definições aparecem em inglês, pois assim estão especificadas na FMA. O mapeamento semântico das regiões da imagem possibilita o uso de qualquer representação do conhecimento, assim poderia ter sido usada a UMLS também. A vantagem da UMLS é que apresenta definições de vários vocabulários biomédicos e em diversas linguagens.

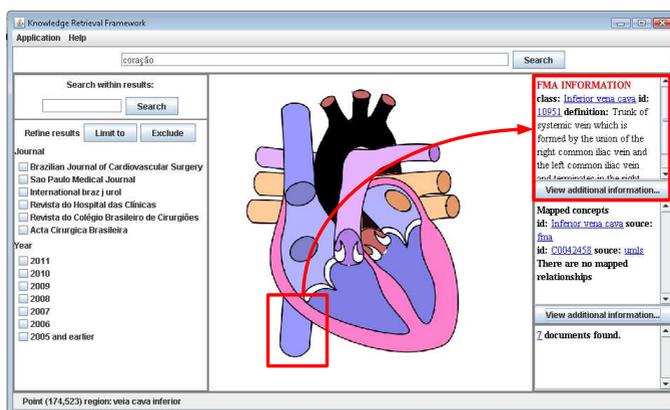


Figura 42 – Informação da região que representa à entidade “veia cava inferior”.

Também usando a imagem recuperada, o usuário pode visualizar quais os conceitos das representações do conhecimento que estão sendo usados para recuperar o conhecimento. A informação apresentada aqui corresponde com o mapeamento semântico feito em cada região da imagem tal como se pode observar na figura 43. No protótipo implementado a informação do mapeamento se apresenta usando a ontologia FMA e o meta-tesauro UMLS.

A quantidade de documentos recuperados a partir dos conceitos mapeados nas regiões da imagem também é apresentada como se pode observar na figura 44. Clicando no enlace do painel se apresenta ao usuário a listagem dos documentos. No protótipo implementado somente foram mapeados o título, as palavras-chaves e o resumo dos artigos.

O outro tipo de visualização implementado no protótipo tenta

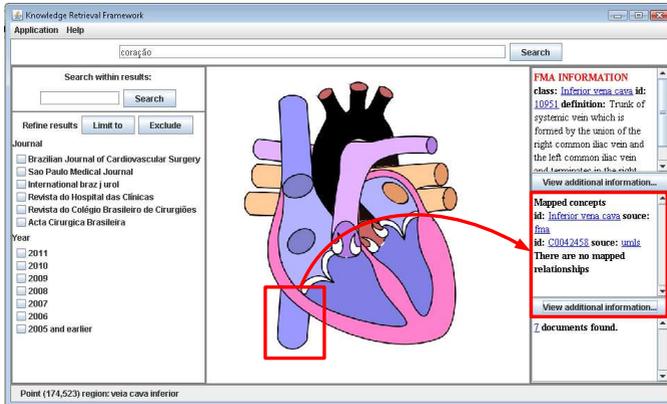


Figura 43 – Informação do mapeamento da região que representa à entidade “veia cava inferior”.

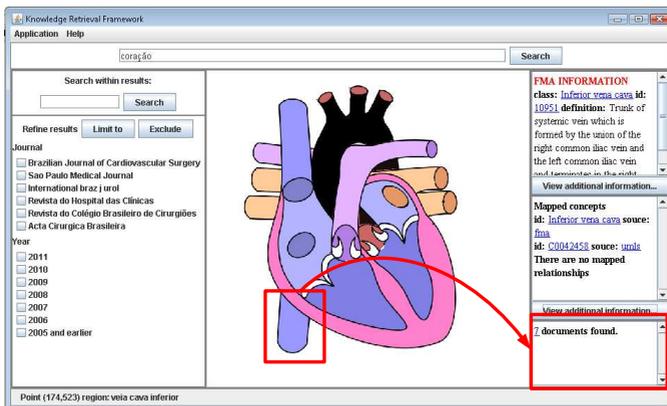


Figura 44 – Quantidade de documentos que mencionam à entidade “veia cava inferior”.

responder à pergunta: *Quantas referências possuem nos repositórios digitais as entidades representadas na imagem?* Na figura 45 apresenta a tela resultante da busca da imagem usando o termo **pulmões**, onde se pode apreciar que agora aparecem números acima das regiões anotadas. Estes números representam a quantidade de documentos que mencionam aos conceitos mapeados em cada região. É outra maneira de visualizar o conhecimento recuperado.

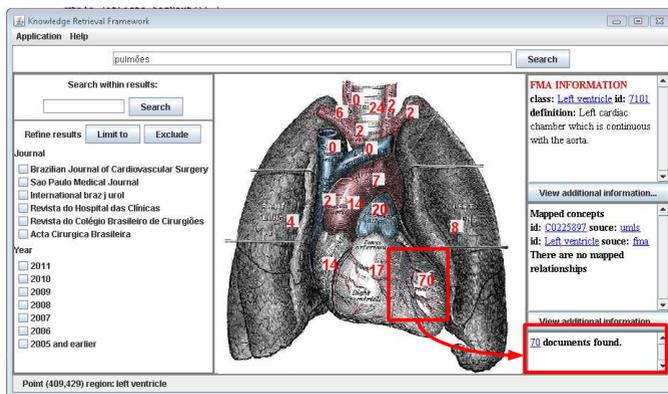


Figura 45 – Tela resultante da busca da imagem pelo termo “pulmões”.

Os usuários podem alterar a informação do mapeamento da imagem para recuperar conhecimento mais específico. Por exemplo, caso o usuário queira recuperar informação relacionada com o ventrículo esquerdo e fumaça de cigarro, poderia alterar o mapeamento para incluir os conceitos da UMLS C0225897 (ventrículo esquerdo) e C0239059 (fumaça de cigarro). A UMLS possui uma rede semântica que pode ser usada para inferir as relações de cada conceito. Na figura 46 pode se observar como alterando o mapeamento, o protótipo retorna agora um documento só que cumpre com a necessidade de informação.

6.2 COMPARAÇÃO DE MODELO PROPOSTO COM OUTRAS ABORDAGENS DE VISUALIZAÇÃO DO CONHECIMENTO

6.2.1 Modelo proposto e o modelo de visualização de Burkhard (2005c)

O modelo de visualização do conhecimento proposto por Burkhard (2005c) (figura 47) é dividido em três partes: um transmissor, um meio e um receptor. Essas três partes estão todas interligadas em um circuito de interação e comunicação.

O modelo descreve processos iterativos inter e intrapessoais: O processo começa com um remetente que pretende transferir alguns dos seus conhecimentos (**conhecimento**) para um destinatário. O modelo

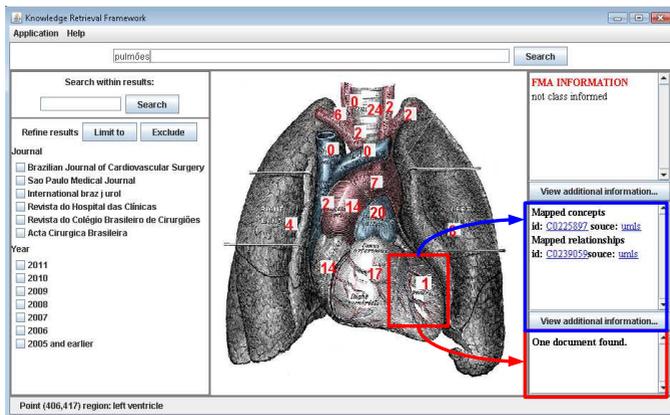


Figura 46 – Filtragem de documentos pelo conceito ‘fumaça de cigarro’.

mental desse conhecimento (**modelo mental do remetente**) é exteriorizada usando várias representações visuais complementares, as quais podem ser divididas em três sub-processos (1, 2, 3 na figura 47) seguindo uma seqüência temporal: Primeiro, o remetente precisa chamar a atenção (1 na figura 47) do destinatário, por exemplo, usando uma imagem provocante. Em segundo lugar, o remetente precisa ilustrar o contexto (2 na figura 47), fornecem uma visão geral (2 na figura 47), e opções para agir (2 na figura 47). Só então o remetente pode apontar para detalhes selecionados (3 na figura 47) que, idealmente, acontece em um diálogo dinâmico com o destinatário (D na figura 47), que re-constrói (C na figura 47) conhecimento semelhantes (**conhecimento**) com essas visualizações complementares e uma própria imagem mental (**modelo mental do destinatário**). Mas devido as diferentes hipóteses, crenças, conhecimentos prévios, inferências e interpretações errôneas podem ocorrer (E na figura 47), o que pode levar a uma falha na formulação do conhecimento novo. Neste processo, o remetente iterativamente refina ou adiciona representações visuais adicionais (F na figura 47), até que o processo de transferência de conhecimentos seja bem sucedido.

Dado que o modelo proposto por Burkhard (2005c) e o proposto neste trabalho se baseiam no *framework* para visualização do conhecimento proposto por Burkhard, existem algumas similitudes entre ambos os modelos. Em particular o processo de visualização em ambos os modelos é muito parecido: primeiro se obtém uma visão global dos dados e depois se concentra em itens de interesse. Também ambos

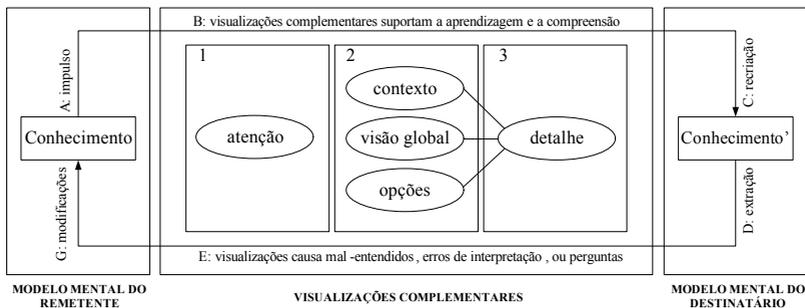


Figura 47 – Modelo para visualização do conhecimento proposto por Burkhard (2005c)

os modelos descrevem as interações dos usuários com representações visuais e podem ser aplicados em qualquer domínio.

O objetivo do modelo de Burkhard (2005c) é recriar modelos mentais nos receptores, no modelo proposto os usuários compartilham o mesmo modelo mental, um modelo mental que faz parte da sua visão do mundo. Os modelos mentais dos remetentes na proposta de Burkhard (2005c) são exteriorizados usando várias representações visuais complementares, no modelo proposto o modelo mental do remetente é representado nas imagens, as quais são conhecidas *a priori*.

A principal diferença com o modelo de Burkhard (2005c) radica no uso de anotações semânticas para descrever as imagens, a proposta de Burkhard (2005c) não usa nenhuma representação do conhecimento para descrever as representações visuais. Por outro lado, o conhecimento que se visualiza no modelo de Burkhard (2005c) se encontra nas visualizações complementares, no modelo proposto, o conhecimento é recuperado a partir dos repositórios de informação. No modelo de Burkhard (2005c) não se recupera o conhecimento a partir de repositórios, este é transferido a partir do modelo mental do remetente.

6.2.2 Modelo proposto e o modelo de visualização RuleViz de Han e Cercone (2000a) e Han e Cercone (2000b)

Em geral, a descoberta de conhecimento em bases de dados é considerada como o processo não trivial de identificação válida, novosa, potencialmente útil e compreensível de padrões nos dados. Basicamente, o processo consiste na coleta dos dados, a criação dos dados

de destino, a limpeza da dados originais, o pré-processamento dos dados, a redução de dados, a mineração de dados, a interpretação dos padrões, e na aplicação do conhecimento. O processo depende de uma interação sistema-humano. É impossível imaginar um processo totalmente automatizado de descoberta de conhecimento, porque isto exige que o sistema possua todo o conhecimento de domínio e ser capaz de reconhecer as intenções do usuário. As ferramentas de visualização podem ser utilizadas para ajudar aos usuários a navegar pelos grandes espaços de informação.

Para construir tais sistemas interativos, Han e Cercone (2000a) propõem RuleViz, um modelo de visualização interativa para visualizar todo o processo de descoberta de conhecimento. É constituída por cinco componentes, como mostrado na figura 48: a i) visualização de dados originais, ii) a redução de dados de forma interativa, iii) o pré-processamento dos dados de forma visual, tais como discretização de atributos numéricos e transformação de dados, iv) a descoberta de padrões como a indução de regras e de construção da árvore de decisão e v) a visualização de padrões.

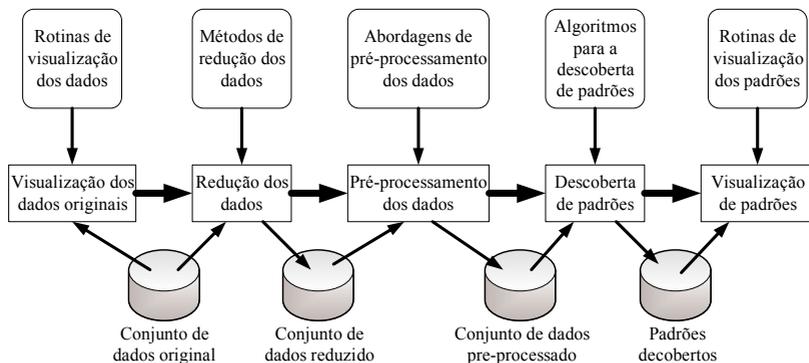


Figura 48 – Modelo para visualização RuleViz proposto por Han e Cercone (2000a) e Han e Cercone (2000b)

Existem várias diferenças no modelo proposto por Han e Cercone (2000a) e o modelo proposto neste trabalho. Uma delas é o domínio de aplicação. O modelo de Han e Cercone (2000a) tem sido concebido para construir sistemas interativos que permitam visualizar os processo da descoberta de conhecimento em banco de dados, os componentes deste modelo está muito ligado as fases do processo de descoberta do conhecimento, pelo que não é aplicável em outro domínio que não seja

a descoberta de conhecimento.

A idéia básica na proposta de Han e Cercone (2000a) é usar técnicas de visualização para visualizar o conjunto de dados originais e obter *insights* sobre a distribuição destes para intuitivamente limitar o domínio dos dados através de interações e assim descobrir as regras a partir dos dados reduzidos e visualizar o conhecimento resultante. A idéia desta tese é bem diferente, ela se baseia na recuperação de conhecimento a partir de repositórios de informação digitais e apresentar este conhecimento usando imagens do mundo real. Na abordagem de Han e Cercone (2000a) são usados diversos métodos e técnicas para a visualização tais como as visualizações geométricas, as visualizações baseadas em ícones, visualizações hierárquicas ou combinações destas. Na nossa proposta não se usam metáforas senão imagens que fazem parte da visão do mundo dos usuários.

6.2.3 Modelo proposto e o modelo de visualização colaborativa proposto por Novak e Wurst (2005)

Novak e Wurst (2005) propõem um modelo (ver figura 49) baseado em artefatos de conhecimento dinâmicos que suportam o acesso dos usuários a espaços de informação de comunidades de uma forma que permite a descoberta de relações. O presuposto deste modelo é que o acesso dos usuários aos repositórios de informação por meio de mapas de conhecimento interativos lhes permite visualizar as informações contidas no repositório a partir de diferentes perspectivas semânticas melhorando assim qualitativamente o processo de formulação do conhecimento que ocorre durante o acesso à informação. Esta abordagem se fundamenta em que o acesso à informação é um processo de criação de significado e aquisição de conhecimento no qual as pessoas através da sua interação com a informação desenvolvem e internalizam novos conhecimentos. Assim, incorporando a capacidade de explorar os repositórios da comunidade não só do ponto de vista pré-definidos da comunidade, mas através das perspectivas dos membros das diferentes comunidades, é uma forma de apoiar a criação e compartilhamento de conhecimento que atravessa as fronteiras das comunidades individuais.

Nesta abordagem, se usa a metáfora mapa do conhecimento como um veículo para descrever a idéia de uma estruturação visual da informação de uma forma, que forneça insights dentro de contextos e relações entre a informação relacionada semanticamente. A diferença fundamental para as abordagens existentes da gestão do conhecimento

e visualização do conhecimento é, que o ponto de partida para a construção dos mapas é o conhecimento pessoal, altamente implícito nos indivíduos e nas comunidades de usuários, ao invés de taxonomias pré-definidas explicitamente ou ontologias. Como o conhecimento das comunidades é muito implícito e socialmente construídos mapas de tal conhecimento não podem ser representações estáticas que codificam o conhecimento. Ao contrário, eles são concebidos como artefatos interativos visuais que podem ser manipulados de forma interativa e discutido pelos membros da comunidade a fim de obter uma compreensão dos diferentes modelos mentais e esquemas interpretativos subjacentes das diferentes comunidades: por exemplo, explorando mapas de diferentes usuários e aplicando-as a situações diferentes.

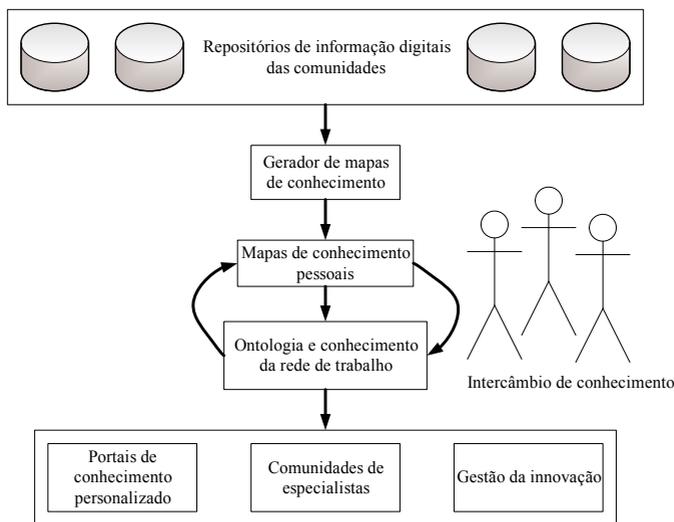


Figura 49 – Modelo de visualização colaborativa proposto por Novak e Wurst (2005)

No modelo proposto por Novak e Wurst (2005), as interações dos usuários se baseiam em mapas de conhecimento os quais estão compostos por dos elementos principais: o mapa do documento e o mapa conceitual. O mapa do documento apresenta as ferramentas de conhecimento do espaço comunitário (por exemplo, documentos, e-mails, posts nos fóruns, descrição de projetos, entre outros), estruturado em grupos de objetos semanticamente relacionados. O mapa conceitual por seu lado, exhibe uma rede de termos que representam grupos de

palavras diferentes usadas em contextos semelhantes e as relações entre eles. Na abordagem desta tese, as interações dos usuários se baseiam em imagens, as quais tem sido enriquecidas com estruturas semânticas a fim de explicitar seu conteúdo.

Na proposta de Novak e Wurst (2005) existe uma necessidade de construção de modelos e ferramentas para visualização de conhecimento dinâmico que forneçam pontos de vista de informações contextualizadas, mostrando suas relações com a perspectiva do conhecimento de uma determinada comunidade. Na nossa abordagem a formulação do conhecimento não é dinâmica e é baseada em representações do conhecimento tais como ontologias, taxonomias e meta-tesauros.

Na abordagem de Novak e Wurst (2005), existe uma visão ontológica que permite ao usuário navegar em um conjunto de documentos com base em uma estrutura pré-semântica. Esta estruturação pode representar pontos de vista pessoais dos membros individuais (chamadas pelo autor de ontologias pessoais) ou opiniões da comunidade, bem como taxonomias que representam domínios de conhecimento formal ou estruturas organizacionais. No modelo proposto na tese, o conhecimento embutido nas imagens por meio dos mapeamentos semânticos não é pessoal, muito pelo contrário estes mapeamentos refletem o conhecimento de um grupo de usuários que compartilham uma mesma visão do mundo, mas estes poderiam ser alterados para recuperar conhecimento mais específico.

6.2.4 Modelo proposto e o modelo para exploração visual de vídeos proposto por Luo et al. (2007)

O modelo proposto por Luo et al. (2007) se concentra principalmente em dois problemas: i) como extrair o conhecimento útil a partir da grande quantidade de informação a partir de bases de dados de notícias em vídeo a grande escala, e ii) como apresentar o conhecimento para os usuários de forma intuitiva. Para resolver estes problemas, o modelo integra vários métodos e técnicas da engenharia do conhecimento tais como a análise semântica de vídeos, a recuperação de informação, a descoberta de conhecimento e a visualização do conhecimento. Além disso, todos os componentes devem ser otimizados para atingir um único objetivo, a exploração intuitiva e inteligente a partir de bancos de dados de vídeo em larga escala.

O modelo proposto pode ser visualizado na figura 50. Primeiro, a interpretação semântica é extraída a partir dos vídeos através de

técnicas de análise semântica de vídeo. Em segundo lugar, a interpretação do conhecimento é extraído através da ponderação da interpretação semântica de acordo com uma medida de interesse. Em terceiro lugar, as técnicas de visualização são adotadas para representar e interpretar o conhecimento de forma intuitiva. Além disso, a interpretação semântica e interpretação do conhecimento pode ser melhorada através do *feedback* do usuário recebidos por uma interface de visualização.

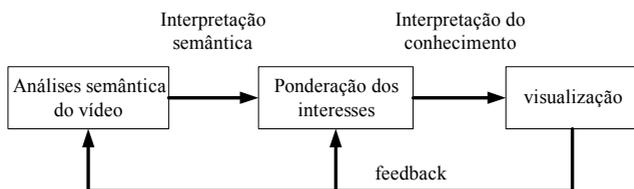


Figura 50 – Modelo para a exploração visual de vídeos proposto por Luo et al. (2007)

Na abordagem de Luo et al. (2007) o tipo de unidade para a interpretação do conhecimento são as palavras-chave dos *closed caption* e os textos associados ao vídeo. Segundo os autores, os usuários são capazes de capturar a idéia principal das reportagens, observando uma seqüência de palavras-chave. Reorganizando estas palavras-chaves com técnicas de visualização adequadas, os usuários são capazes de capturar a semântica subjacente e o conhecimento oculto intuitivamente. Esta abordagem é similar à proposta desta tese, em ambas as propostas o modelo precisa entender o que está sendo representado, porém no modelo proposta nesta tese, a interpretação do conhecimento é feita usando mapeamentos semânticos usando representações do conhecimento formais como ontologias e meta-tesauros.

As visualizações na proposta de Luo et al. (2007) visam apresentar ao usuários itens de conhecimento para que os vídeos possam ser explorados de uma forma eficiente e intuitiva, usando diversas metáforas visuais como as visualizações radiais. O foco principal desta proposta é a exploração de vídeos e não a recuperação do conhecimento.

6.2.5 Modelo proposto e o sistema de recomendação do conhecimento proposto por Li e Lu (2010)

Com o objetivo de ajudar aos usuários a adquirir e visualizar de forma efetiva conhecimento a partir de repositórios de informação massivos, distribuídos e heterogêneos, Li e Lu (2010) propõem um sistema para recomendação de conhecimento baseado em mapas de tópicos inteligentes. Este inclui a organização do conhecimento, a recomendação do conhecimento e a visualização do conhecimento (ver figura 51).

Quando o cliente Web envia uma solicitação ao portal do conhecimento, o componente recuperação do conhecimento adquire este a partir da base de dados dos mapas de tópicos inteligentes. Dado que os registros históricos dos usuários são salvos, o serviço de recomendação do conhecimento pode fornecer recomendações ao componente de visualização do conhecimento. O sistema de recomendação de conhecimento inclui a descoberta de recursos, a predição das tendências e a navegação do conhecimento.

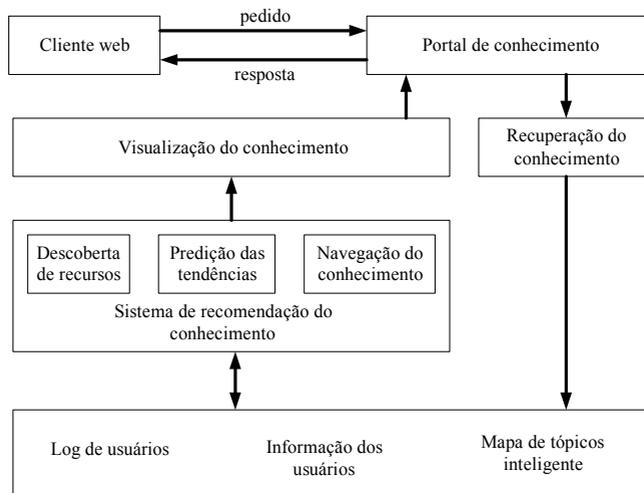


Figura 51 – Sistema de recomendação do conhecimento proposto por Li e Lu (2010)

A principal diferença da proposta de Li e Lu (2010) e a da tese encontra-se nos objetivos, os sistema proposto por Li e Lu (2010) procura por conhecimento de interesse do usuário, evitando assim que o usuário tenha que navegar pelos repositórios grandes repositórios de

informação. O interesse do usuário é obtido a partir do histórico das buscas feitas pelos usuários assim como pelos logs gerados pelo sistema para cada usuário. No modelo proposta nesta tese não se modela a princípio o interesse de cada usuário e sim de grupos de compartilham a mesma visão do mundo. Este interesse se traduz nos mapeamentos semânticos, porém os usuários encontram-se habilitados para refinar estes mapeamentos e assim traduzir seus interesses pessoais, mas sempre sob uma base de interesse grupal.

A abordagem de Li e Lu (2010) baseia-se em mapas de tópicos inteligentes, os quais encontram-se organizados em camadas tais como: *clusters*, tópicos, elementos de conhecimento e recursos. Estes mapas de tópicos inteligentes definem a estrutura do conhecimento para o sistema tendo que ser obtidos para cada domínio de aplicação. Na proposta da tese, a diferença desta proposta optou por se usar conhecimento explicitamente definidos em representações do conhecimento, tais como ontologias, taxonomias e meta-tesauros.

6.2.6 Modelo proposto e o *framework* para a visualização do conhecimento baseada em *dashboards* proposto por Mohd, Embong e Zain (2010)

Mohd, Embong e Zain (2010) propõem uma abordagem para a visualização do conhecimento adotando técnicas de visualização baseada em grafos e incorporando o conceito de *dashboards* para instituições de ensino superior. Esta abordagem enfatiza em dois aspectos: a visualização do conhecimento e a interação humano-computador. A visualização do conhecimento ajuda aos usuários a analisar as várias características dos estudantes e professores após um processo de clusterização, por outro lado, a interação permite a transferência de conhecimento e o uso das capacidades de percepção humanas incrementando assim a inteligência do sistema. A visualização do conhecimento é reforçada através dos *dashboard* que fornece padrões de conhecimento.

O sistema começa com o processo de clusterização para agrupar a informação existente no banco de dados. Ao mesmo tempo, os inputs feitos no indicador de performance gerarão os respectivos *dashboard*. Para ambos os processos, a interação com os especialistas no domínio melhora a inteligência do sistema. As saídas dos processos de clusterização e processamento dos *dashboard*, são processados pelo componente de visualização com o objetivo de produzir como resultado final um *dashboard* interativo. A arquitetura deste sistema pode ser

apreciada na figura 52.

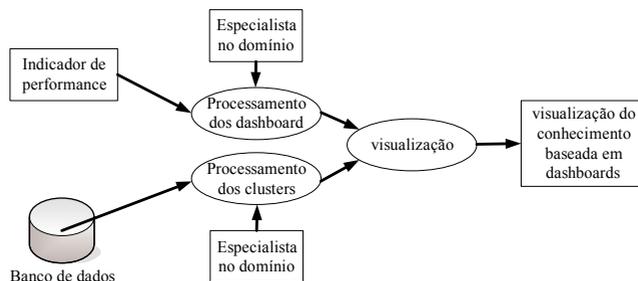


Figura 52 – *Framework* para a visualização do conhecimento baseada em *dashboards* proposto por Mohd, Embong e Zain (2010)

A principal característica da abordagem de Mohd, Embong e Zain (2010) é a interação de especialistas do domínio que incorporam a inteligência ao sistema, na abordagem da tese todo o conhecimento é explicitado nos mapeamentos semânticos os quais vinculam a conceitos de diversas representações do conhecimento.

6.3 ANÁLISE DO MODELO POR MEIO DE HEURÍSTICAS

Na revisão sistemática, encontraram-se dois estudos que fazem uma meta-análise com o intuito de aumentar entendimento do uso de diferentes conjuntos de heurísticas para avaliação de visualização de informações. Destes estudos, Zuk et al. (2006) e Forsell e Johansson (2010) agrupam as heurísticas em categorias que abrangem uma grande variedade de perspectivas sobre usabilidade relevantes para os sistemas de visualização da informação visualização da informação. Baseado nestas categorias, foi analisado o modelo e o protótipo, os resultados desta análise se apresenta na seqüência.

6.3.1 Modelo proposto e as heurísticas perceptivas e cognitivas de Zuk e Carpendale (2006)

Zuk e Carpendale (2006) propõem um conjunto de heurísticas para a aplicação das teorias perceptivas e cognitivas para a análise das visualizações. Estas heurísticas se baseiam nas variáveis de visualização de Bertin (1983), nos princípios de Tufte e Howard (2001) para a ex-

celência em gráficos e no livro para visualização da informação de Ware (2000). Na seqüência, apresentam-se as heurísticas compiladas por Zuk e Carpendale (2006) e se analisa cada uma delas em relação ao modelo e protótipo proposto.

- **Assegurar-se que a variável visual tenha o tamanho suficiente (BERTIN, 1983; WARE, 2000; ZUK; CARPENDALE, 2006):** Na implementação do protótipo, as imagens semânticas são apresentadas no painel central, ocupando aproximadamente o 50% do tamanho da tela. As imagens representam no nosso modelo as variáveis visuais, e estas são destacadas no protótipo com um tamanho adequado como se pode apreciar nas figuras 41, 42, 43, 44, 45 e 46.
- **Não espere que uma ordem de leitura a partir da cor (BERTIN, 1983; WARE, 2000; ZUK; CARPENDALE, 2006):** A variável cor não tem ordem implícita, mas é seletiva e associativa. No modelo é usada a cor para associar a quantidade de documentos recuperados a partir de uma região da imagem, mudando a coloração original da região como se pode apreciar na figura 41. No modelo não se espera uma ordem implícita de leitura a partir da cor.
- **A percepção da cor varia com o tamanho do item (BERTIN, 1983; WARE, 2000; ZUK; CARPENDALE, 2006):** No modelo são usadas imagens que representam entidades do mundo real. A cor, como dito anteriormente, é usada para associar a quantidade de documentos recuperados a partir de uma região. Devido a que a percepção da cor varia com o tamanho do item, no prototipo implementado, somente é mudada a cor da região com maior quantidade de documentos associados. Este tipo de mudança de cores não pode ser usada em todas as regiões pois não teria a mesma percepção e confundiria aos usuários.
- **O contraste local afeta a percepção das cores (ZUK; CARPENDALE, 2006; WARE, 2000):** Devido a esta heurística, a cor usada para alterar a cor da região com maior quantidade de documentos, é de uma contrasta com as cores originais da imagem, isto justamente para alterar a percepção do usuário com o intuito de que rapidamente a reconheçam esta região.
- **Considerar as pessoas com daltonismo (ZUK; CARPENDALE, 2006):** O daltonismo é uma perturbação da percepção

visual onde se vê cores caracterizadas pela incapacidade de diferenciar todas ou algumas cores, manifestando-se muitas vezes pela dificuldade em distinguir o verde do vermelho. No modelo proposto não tem sido considerada esta situação, mas poderia ter sido tratada caracterizando as preferências de visualização do usuário, por exemplo mudando a visualização por cores pelo uso da intermitência.

- **Os benefícios pré-atentivos aumentam com o campo de visão (ZUK; CARPENDALE, 2006; WARE, 2000):** A característica pré-atentiva usada foi a cor. Esta tem sido usada sobre a mesma imagem, no campo de visão do usuário.
- **Preservar os dados para gráficar a dimensionalidade (BERTIN, 1983; TUFTE; HOWARD, 2001; ZUK; CARPENDALE, 2006):** Os mapeamentos semânticos feitos em cada região da imagem permitem que os dados a serem apresentados sejam recuperados a partir dos repositórios de documentos sempre que seja necessário.
- **Coloque mais dados em menos espaço (ZUK; CARPENDALE, 2006; TUFTE; HOWARD, 2001):** Todos os dados são apresentados na imagem no painel central no mesmo espaço.
- **Remova o elementos estranhos (TUFTE; HOWARD, 2001; ZUK; CARPENDALE, 2006):** O protótipo implementado tem como elemento principal a imagem a partir da qual se recuperará o conhecimento dos repositórios digitais. A imagem é colocada no painel central da tela principal e nenhum elemento adicional é incluído na tela, visando diminuir a carga de trabalho cognitiva do usuário.
- **Considere as leis da Gestalt (ZUK; CARPENDALE, 2006):** Devido a que se usam imagens do mundo real e não metáforas visuais, não foi necessário considerar as leis da Gestalt.
- **Fornecer vários níveis de detalhe (TUFTE; HOWARD, 2001; WARE, 2000; ZUK; CARPENDALE, 2006):** A imagem no painel central é a variável usada para apresentar a visão global dos dados recuperados a partir dos repositórios. Selecionando uma região da imagem e estando esta mapeada semanticamente, o protótipo pode apresentar a descrição do conceito, também é possível recuperar a lista dos documentos associados a cada imagem e desta lista visualizar um documento em específico.

- **Integrar texto sempre que relevante (TUFTE; HOWARD, 2001; WARE, 2000; ZUK; CARPENDALE, 2006):** No protótipo se integra o texto na imagem quando é preciso conhecer a quantidade de documentos associados por cada região como se pode apreciar nas figuras 45 e 46.

6.3.2 Modelo proposto e o mantra da visualização da informação de Shneiderman (1996)

Um bom ponto de partida para a concepção de interfaces gráficas de usuário é o mantra da visualização da informação que se resume em: apresentar primeiro a visão global dos dados, fazer depois a filtragem dos dados e finalmente recuperar os detalhes sob demanda. Na seqüência, apresentam-se as heurísticas propostas por Shneiderman (1996) e se analisa cada uma delas em relação ao modelo e protótipo proposto.

- **Visão global (SHNEIDERMAN, 1996; CRAFT; CAIRNS, 2005):** A idéia desta heurística é que a variável visual apresente ao usuário uma visão geral dos dados. Dado que no modelo a variável visual usada são as imagens semânticas, a visão global é feita nestas estruturas. Por exemplo, para saber quantos documentos associados possui cada entidade que compõe a imagem é apresentada a quantidade, como texto, encima de cada região como se pode apreciar nas figuras 45 e 46.
- **Zoom e filtragem (SHNEIDERMAN, 1996):** A idéia desta heurística é que os usuários se concentrem nos itens do seu interesse. Os usuários normalmente têm interesse em alguma parte de uma coleção, e eles precisam de ferramentas que os ajudem nesta tarefa. Por outro lado, devem filtrar os itens não relevantes. De forma geral se deve permitir ao usuário controlar o conteúdo da visualização. O modelo permite a filtragem dos dados de várias formas: i) utilizando os metadados dos repositórios de documentos, ii) utilizando os conceitos dos artefatos de conhecimento (filtragem semântica), e iii) utilizando filtragem textual.
- **Detalhes sob demanda (SHNEIDERMAN, 1996; CRAFT; CAIRNS, 2005):** A idéia desta heurística é que ao selecionar um item seja possível obter os detalhes quando necessário. O modelo permite isto através da seleção das regiões da imagem. Ao selecionar uma região da imagem, além de apresentar ao usuário

informações sobre o conceito, também é apresentada a quantidade de documentos relacionados com os conceitos mapeados em cada região, sendo possível visualizar a lista de todos os documentos do repositório relacionados a esse conceito e podendo recuperar um documento em específico.

- **Relacionamentos (SHNEIDERMAN, 1996; CRAFT; CAIRNS, 2005):** A idéia desta heurística é que o usuário possa ver as relações entre os itens. No modelo estas relações são facilitadas pelos artefatos de conhecimento que permite que possam ser feitas inferências sobre os conceitos.
- **Extração (SHNEIDERMAN, 1996; CRAFT; CAIRNS, 2005):** A idéia desta heurística é que o usuário possa extrair sub-coleções e os parâmetros usados nas consultas. No modelo proposto não tem sido considerada esta situação.
- **História (SHNEIDERMAN, 1996; CRAFT; CAIRNS, 2005):** A exploração da informação é inerentemente um processo com muitas etapas, mantendo assim o histórico das ações e permitindo que os usuários refaçam seus passos é um importante apoio a este processo. No modelo proposto não tem sido considerada esta situação.

6.3.3 Modelo proposto e critério de Freitas et al. (2002)

O trabalho de Freitas et al. (2002) apresenta critérios para avaliar técnicas de visualização, em relação a representações visuais e mecanismos de interação. Estes critérios têm sido propostos para avaliar métodos e técnicas para apresentar informação hierarquizada. Dado que o modelo não apresenta informação hierarquizada, não tem sido considerados estes critérios para analisar o modelo.

6.3.4 Modelo proposto e as dez heurísticas de usabilidade de Nielsen e Mack (1994)

Nielsen e Mack (1994) propõem dez heurísticas para avaliar a usabilidade de um sistema. Na seqüência, apresentam-se as heurísticas compiladas propostas por Nielsen e Mack (1994) e se analisa cada uma delas em relação ao modelo e protótipo proposto.

- **Correspondência entre o sistema e o mundo real:** O sistema deve falar a linguagem do usuário, com palavras, frases e conceitos familiares ao usuário, ao invés de termos orientados ao sistema. Deve seguir as convenções do mundo real, de forma tal que a informação apareça em uma ordem natural e lógica. No modelo, as imagens têm a característica de ser representações do mundo real, representações que são conhecidas pelo grupo alvo.
- **Controle do usuário e liberdade:** Os usuários freqüentemente escolhem funções do sistema por engano e vão precisar de uma “saída de emergência” claramente marcada para sair do estado indesejado sem ter que passar por um extenso diálogo com o sistema. No protótipo implementado, as opções do usuário são basicamente de filtragem e recuperação de documentos sob demanda, mas caso o usuário precise desfazer as opções de filtragem, pode voltar ao estado original (antes de fazer as filtrações). Isto é facilitado pelos mapeamentos semânticos feitos na imagem, sempre será possível recuperar o conhecimento a partir das regiões da imagem.
- **Consistência e padrões:** Os usuários não precisam adivinhar que diferentes palavras, situações ou ações significam a mesma coisa. O protótipo implementado possui como variável de visualização as imagens do mundo real as quais têm sido mapeadas previamente. As opções de recuperação de conhecimento são consistentes em todo o protótipo e são feitas selecionando a região da imagem.
- **Reconhecimento ao invés de recordação:** Minimizar a carga de memória do usuário fazendo os objetos de decisões, ações e opções visíveis. O usuário não deve ter que lembrar informação de uma parte para outra. As instruções para a utilização do sistema devem estar visíveis e ser facilmente recuperáveis quando necessárias. No protótipo desenvolvido, o usuário dialoga com o sistema por meio das imagens, desta forma se evita acionar a memória de longo prazo do usuário para, por exemplo, lembrar dos termos da consulta.
- **Flexibilidade e eficiência de uso:** Os aceleradores, despercebidos pelos usuários iniciantes, podem muitas vezes acelerar a interação do usuário experiente de tal forma que o sistema possa atender tanto aos usuários inexperientes quanto aos experientes.

Permitem aos utilizadores freqüentes personalizar ações. Esta heurística não tem sido considerada ao desenvolver o protótipo.

- **Estética e *design* minimalista:** As telas não devem conter informação irrelevante ou desnecessária. Os “diálogos” do sistema precisam ser simples, diretos e naturais, presentes nos momentos em que são necessários. No modelo proposto, a comunicação do usuário com o sistema se dá através das imagens, e sob as regiões destas imagens somente é colocada informação relevante, seja alterando a cor da região ou colocando as quantidades dos documentos em forma de texto. Todo isto com o intuito de usar o processamento pré-atentivo do cérebro e evitar a carga de trabalho cognitivo.
- **Prevenção de erros:** Ainda melhor do que boas mensagens de erro é um projeto cuidadoso que impede que um problema ocorra em primeiro lugar. Devem ser eliminadas as condições propensas aos erros ou verificá-los e apresentar ao usuário uma opção de confirmação antes de se comprometerem com a ação. Existem situações que podem fazer que o sistema não funcione (*i.e.*, queda do servidor de banco de dados, falhas no motor de recuperação de informação, erro ao acessar o banco de dados das imagens, etc.), mas estas situações são devidamente tratadas no protótipo desenvolvido.
- **Ajudar os usuários a reconhecer, diagnosticar e se recuperar dos erros:** As mensagens de erro devem ser expressas em linguagem clara (sem códigos), indicar com precisão o problema e construtivamente sugerir uma solução. As mensagens de erro tem sido elaborados tendo em consideração esta heurística.
- **Ajuda e documentação:** Mesmo que seja melhor se o sistema pode ser usado sem documentação, pode ser necessário fornecer ajuda aos usuários em relação ao seu uso. Tais informações devem ser fáceis de pesquisar, focada na tarefa do usuário, e não ser muito grande. O protótipo desenvolvido não possui ajuda online.
- **Visibilidade do status do sistema:** O sistema deve sempre manter os usuários informados sobre o que está acontecendo, através de *feedback* apropriado em tempo razoável. Esta heurística não tem sido considerada ao desenvolver o protótipo.

7 CONCLUSÕES E TRABALHOS FUTUROS

7.1 PUBLICAÇÕES RELACIONADAS AO TRABALHO DA TESE

No decorrer trabalho e como parte dos requisitos do programa, elaboraram-se uma série de publicações relacionadas com o assunto do projeto de pesquisa.

A idéia inicial do projeto era a criação de um ambiente para visualizar o conhecimento armazenado na memória organizacional. Como parte desta fase dois artigos foram elaborados, um deles fazendo uma revisão da aplicação da engenharia do conhecimento na área de memória organizacional (MELGAR; BEPLER; PACHECO, 2011) e o outro propondo uma proposta de um sistema para a visualização do conhecimento baseado em métodos e técnicas da engenharia do conhecimento aplicado na área da saúde (MELGAR; PACHECO, 2010a).

Posteriormente, o foco mudou para a visualização do conhecimento a partir de repositórios de informação. Nesta fase, elaborou-se um artigo contendo a arquitetura inicial proposta baseada na ontologia FMA (MELGAR; PACHECO, 2010c). Dado que o protótipo foi desenvolvido usando como caso de aplicação o domínio da anatomia, alguns artigos foram elaborados para relatar os primeiros resultados da pesquisa, incluindo as especificação dos mapeamentos semânticos nas imagens (MELGAR; PACHECO, 2010d, 2010b), o uso da recuperação da informação para recuperar documentos a partir das imagens semânticas (MELGAR; BEPLER; PACHECO, 2010b) e um *framework* para a recuperação de conhecimento anatômico usando a ontologia FMA e o metatesauro UMLS (MELGAR; BEPLER; PACHECO, 2010a).

Finalmente, foi elaborado o modelo para visualização proposto nesta tese, a primeira versão deste modelo trabalhava com o conceito de arquétipos visuais (MELGAR; BEPLER; PACHECO, 2011), mas na versão final do modelo se optou por mudar para o nome de imagens semânticas (MELGAR et al., 2011).

7.2 CONCLUSÕES

O objetivo geral desta tese é desenvolver um modelo que, usando estruturas semânticas que representam imagens do mundo real, possibilite a visualização do conhecimento armazenado em repositórios de documentos digitais. Recorreu-se i) à visualização da informação para

compreender como os seres humanos processam as imagens no nosso cérebro, ii) à visualização do conhecimento para fornecer o quadro conceitual para a transferência de conhecimento por meio de imagens do mundo real, iii) à teoria da carga cognitiva para compreender como a apresentação da informação permite aos seres humanos a construção de modelos mentais que facilitam o seu aprendizagem, iv) à semiótica para entender as estruturas de produção de significados e a troca de mensagens e v) à web semântica que forneceu os mecanismos para enriquecer as imagens com significado entendível pelos computadores.

O modelo proposto visa explicar como os usuários interagem com as imagens com o objetivo de visualizar e recuperar conhecimento a partir de repositórios de documentos digitais. Baseados neste modelo os engenheiros do conhecimento podem desenvolver sistemas para serem usados em vários domínios do conhecimento (*i.e.*, saúde, mecânica, aeronáutica, etc.) e usando diversos tipos de repositórios de documentos (*i.e.*, repositórios de artigos científicos, memória organizacional, repositórios de competências, documentos internos, etc.). O modelo não define os tipos de documentos armazenados nos repositórios, a princípio qualquer tipo de documento poderá ser armazenado e portanto recuperado.

Os mecanismos de integração (*i.e.*, mapeamentos semânticos) permitem ao modelo compreender tanto as imagens quanto os documentos dos repositórios, de forma tal que ao selecionar uma região de uma imagem, o modelo saiba qual o conceito que foi visualizado. Usando este conceito, o modelo é capaz de recuperar, a partir dos repositórios de documentos, as informações associadas ao conceito.

No modelo os mecanismos de integração são implementados usando representações do conhecimento do domínio (*i.e.*, ontologias, taxonomias, meta-tesauros, etc.). Isto oferece algumas vantagens adicionais ao modelo como, por exemplo, a busca conceitual independente da linguagem do documento e a possibilidade de inferir informações adicionais. O modelo permite também que possam ser usadas em paralelo várias representações de conhecimento para realizar os mapeamentos. Desta forma o modelo pode oferecer aos usuários uma maior quantidade de dimensões para a filtragem dos dados.

O modelo usa as imagens semânticas como uma nova forma para explorar o conhecimento residente nos repositórios digitais. Neste modelo o usuário tem que selecionar a imagem a partir da qual será visualizado o conhecimento. Usando a imagem, o usuário não precisa ingressar os termos para realizar as consultas nos repositórios, pois como o modelo é capacitado para compreender quais as entidades representadas

na imagem, usa os conceitos relacionados a estas entidades e monta as respectivas consultas, ajudando assim a traduzir as necessidades de informação dos usuários.

Além disso, como o modelo pode interpretar a estrutura da imagem, este conhecimento é usado para alterar as regiões das imagens de acordo com os resultados das buscas. A idéia é que os usuários possam identificar rapidamente as regiões nas imagens onde possuem documentos relacionados. Isto pode ser feito, por exemplo, alterando a coloração das regiões onde se concentra a maior quantidade dos documentos ou incluindo a quantidade dos documentos encontrados relacionados às regiões.

O modelo é composto por cinco componentes: os usuários, as imagens semânticas, os repositórios de documentos, o repositórios de conhecimento, e o componente de visualização. De forma geral o modelo descreve as interações dos usuários com as imagens semânticas a fim de satisfazer uma necessidade de informação. A visão de mundo é usada para criar as imagens semânticas com o objetivo que os usuários compreendam rapidamente a informação visual representados neles e possam relacionar conhecimentos prévios. Para que o modelo compreenda os conceitos representados nas imagens e os mencionados nos documentos foram usados mapeamentos semânticos.

As interações dos usuários são feitas através de um componente de visualização, é neste componente que são acionados os mecanismos do modelo: a visualização das imagens permite ao usuário selecionar e visualizar uma imagem, a visualização do conhecimento apresentará na própria imagem os resultados das buscas dos conceitos associados à imagem, o refinamento permite ao usuário concentra-se na informação do seu interesse, e a visualização dos documentos permite que o usuário visualize uma determinada peça da informação.

No modelo tanto as imagens quanto os repositórios de documentos têm sido enriquecidos com conteúdo semântico possibilitando a integração destes dois componentes. O conteúdo semântico permite ao protótipo “entender” quais os conceitos representados tanto nas imagens quanto nos documentos facilitando assim a busca conceitual. Uma das vantagens disto é a recuperação de conhecimento de forma independente à linguagem da escrita dos documentos. Devido à este “entendimento”, o protótipo é capaz de recuperar os documentos relacionados aos conceitos representados nas imagens auxiliando ao usuário no processo de busca. Este mecanismo de recuperação pode ser observado no protótipo, para recuperar documentos o usuário só precisa selecionar uma região da imagem sem necessidade de ingressar os termos das

buscas como nos sistemas de recuperação tradicionais. Entretanto, o usuário deve ingressar termos para procurar imagens. Uma vez selecionado a imagem o modelo “entende” quais as informações necessita recuperar usando para tal os mapeamentos semânticos.

A idéia do modelo é que os resultados das buscas sejam apresentados de forma global no próprio arquétipo alterando a coloração das regiões onde se concentra a maior quantidade de documentos. Este comportamento do modelo é consistente com as tarefas de visualização definidas por Shneiderman (1996) (*i.e.*, primeiro apresentar a visão global dos dados, depois concentrar-se em itens de interesse e filtrar itens irrelevantes e finalmente fornecer a informação sob demanda).

O motivo para usar imagens contendo representações do mundo real se baseou por um lado nas habilidades que dos seres humanos em processar imagens rapidamente e por outro lado na facilidade dos indivíduos para relacionar conhecimentos prévios associados a uma imagem já conhecida. Ao ver uma imagem que pertence a sua visão do mundo, as pessoas sabem o que ela representa apesar de não se lembrar dos nomes dos conceitos nela representados.

Uma das vantagens que oferece o modelo é a facilidade para recuperação de documentos usando apenas regiões das imagens. Nos sistemas de recuperação tradicionais, os usuários traduzem suas necessidades de informação em termos de buscas, estes retornam uma lista de itens que correspondem aos documentos mais relevantes segundo os termos informados. No modelo proposto os usuários traduzem suas necessidades de informação também em termos, mas estes termos não são usados para procurar diretamente os documentos e sim para procurar as imagens que será usadas no processo de visualização.

Outra vantagem do modelo é a integração dos componentes por meio dos mapeamentos semânticos. As imagens e os diferentes repositórios de informação encontram-se todos integrados por meio de informação semântica, isto permite que uma mesma imagem possa ser usada para visualizar documentos de diferentes repositórios. Por exemplo, a imagem do coração apresentado nas seções anteriores, pode ser usado para recuperar artigos científicos, competências organizacionais, projetos, estudos clínicos, indicadores, imagens médicas, entre outros. Esta integração semântica também permite que o modelo possa inferir novas informações sobre um determinado conceito da imagem. Por exemplo, ao projetar o coração, é possível recuperar informações relacionadas a conceitos que não são definidos de forma explícita na imagem, mas que podem ser inferidos, como por exemplo, certas doenças. No coração ao selecionar a aorta, por exemplo, e usando as relações UMLS

poderiam ser recuperados os documentos relacionados ao aneurisma, uma doença que afeta a aorta.

7.3 TRABALHOS FUTUROS

O modelo apresentado baseia-se no uso de imagens do mundo real, as quais têm sido mapeadas usando uma visão do mundo comparilhada, razão pela qual todos os usuários usam os mesmos mapeamentos semânticos para a recuperação do conhecimento. Uma interessante alteração no modelo seria a inclusão de mecanismos que permitam aos usuários informar seus interesses de forma explícita sobre certos conceitos e assim o modelo “saiba” quais os conhecimentos que o usuário tem interesse. Os usuários ao selecionar as regiões da imagem só recuperariam o conhecimento do seu interesse. Por outro lado estes interesses poderiam ser usados para comunicar ao usuário a chegada de novo conhecimento “interessante”, para isso o modelo deveria poder inferir o que é um conhecimento novo para cada usuário. Muitas vezes os usuários não podem definir seus interesses de forma explícita, isto devido a que em muitas vezes não são cientes dos seus interesses ou não possuem tempo para fazê-lo. Para atender esta problemática, uma alteração que poderia ser feita no modelo é a inclusão de mecanismos que permitam inferir os interesses dos usuários baseados nos históricos das buscas.

No modelo, a visualização do conhecimento se expressa por meio de alteração da cor nas regiões da imagem para comunicar a concentração de documentos que se relacionam com a entidade selecionada, ou apresentando números acima das regiões que indicam quantos documentos se encontram nos repositórios de informação. No modelo não existem mecanismos que permitem agrupar os documentos recuperados por cada região com o objetivo de identificar quais os documentos relacionados. Entende-se que isto poderia ser solucionado usando visualizações complementares tipo árvores.

Em relação à ferramenta, na implementação do modelo somente foram consideradas imagens em duas dimensões, mas os conceitos do modelo também funcionam com imagens em três dimensões. Sugere-se uma implementação de uma ferramenta baseada em *voxels*. Um *voxel* (*VOLumetric piXEL*) representa um valor em uma matriz regular em um espaço tridimensional. Este elemento é análogo ao *pixel* que é usado para representar imagens em duas dimensões (ELVINS, 1992; GALLAGHER, 1995).

Mapeando as imagens a outras ontologias, é possível aplicar o

modelo para suportar a visualização do conhecimento em diversas áreas. A gestão por competências poderia ser uma delas. A gestão por competências pode ser definida como a maneira pela qual as organizações gerenciam as competências da corporação, os grupos e os indivíduos. Esta organiza-se em quatro tipos de processo: identificação de competências i), ii) avaliação de competências), a aquisição de competências iii) e iv) uso do conhecimento sobre as competências (BERIO; HARZALLAH, 2005, 2007). Para representar competências, vários modelos e linguagens foram propostos (VASCONCELOS et al., 2000; HARZALLAH; BERIO, 2004; HARZALLAH; BERIO; VERNADAT, 2006; BERIO; HARZALLAH, 2007; TARASSOV; SANDKUHL; HENOCH, Feb. 12-16, 2006; PÉPIOT et al., 2007) usando tanto técnicas da engenharia de *software* como da engenharia do conhecimento. Adicionando ao modelo uma ontologia que represente o domínio da gestão por competências e mapeando esta ontologia nas imagens, será possível a visualização de conhecimento relacionado a determinadas competências. Seguindo com o exemplo do coração e mapeando este a uma ontologia que represente as competências dos indivíduos, este poderia ser usado para identificar por exemplo, se a organização tem competências relacionadas com a aorta. Poderia-se filtrar os tipos de competências que o usuário precisar (*i.e.*, cirurgia, diagnósticos, interesses, etc).

Outra aplicação poderia ser no campo da memória organizacional. De forma geral a memória organizacional pode ser definida como um sistema capaz de armazenar as coisas percebidas, experimentadas ou vividas para além da duração da ocorrência atual, e permitir recuperá-las em um momento posterior (LEHNER; MAIER, 2000, p. 283). É o conhecimento de como fazer as coisas, a forma de abordar os problemas e questões, como tratamos uns aos outros (JACKSON, 2007, p. 90). A memória organizacional está preocupada com a possibilidade de reutilizar a experiência de uma organização (ACKERMAN; HADVERSON, 2000, p. 59). Neste campo também vários modelos e meta-modelos têm sido proposto para representar a memória organizacional (RAMESH, 1997; ABECKER et al., 1998; NEVO; WAND, 2005; JACKSON, 2007). Incorporando ao modelo uma ontologia que representa a memória organizacional, este poderia ser utilizado para visualizar conhecimento relacionado aos retentores de conhecimento ou as tarefas que fazem uso deste conhecimento.

O trabalho proposto traz como elemento principal as imagens do mundo real que fazem parte da visão de mundo dos usuários. Uma questão interessante e que não foi abordada neste trabalho é: Como a visualização do conhecimento e as imagens semânticas podem apoiar

na disseminação do conhecimento para grupos de usuários que não compartilham a mesma visão de mundo? Uma proposta a esta resposta poderia estar na área da hipermídia adaptativa.

Há a necessidade de efetuar experimentos mais extensos a fim de medir os benefícios da visualização do conhecimento apresentado no modelo da tese. O objetivo aqui será avaliar os possíveis benefícios no contexto da utilização do modelo de uso cotidiano no dia-a-dia do usuários.

REFERÊNCIAS

- ABECKER, A. et al. Toward a technology for organizational memories. *IEEE Intelligent Systems*, IEEE Educational Activities Department, Piscataway, NJ, USA, v. 13, n. 3, p. 40–48, 1998. ISSN 1541-1672.
- ACKERMAN, M. S.; HADVERSON, C. A. Reexamining organizational memory. *Commun. ACM*, v. 43, n. 1, p. 58–64, 2000.
- AGOSTI, M.; FERRO, N. A formal model of annotations of digital content. *ACM Trans. Inf. Syst.*, v. 26, n. 1, p. 3, 2007.
- AHLBERG, C.; SHNEIDERMAN, B. Visual information seeking: tight coupling of dynamic query filters with starfield displays. In: *Proceedings of the SIGCHI conference on Human factors in computing systems: celebrating interdependence*. Boston, Massachusetts, United States: ACM, 1994. p. 313–317.
- AHLBERG, C.; WISTRAND, E. Ivey: an information visualization and exploration environment. In: *Proceedings of Information Visualization*. [S.l.: s.n.], 1995. p. 66–73.
- AKSOY, D. Information source selection for resource constrained environments. *SIGMOD Rec.*, v. 34, n. 4, p. 15–20, 2005.
- ALAVI, M.; LEIDNER, D. Review: Knowledge management and knowledge management systems: Conceptual foundations and research issues. *MIS Quarterly*, v. 25, n. 1, p. 107–136, 2001.
- AMAR, R.; STASKO, J. A knowledge task-based framework for design and evaluation of information visualizations. In: . [S.l.: s.n.], 2004. p. 143–150. ISSN 1522-404X.
- ANDREWS, K.; HEIDEGGER, H. Information slices: Visualising and exploring large hierarchies using cascading, semi-circular discs. In: *Proceedings of IEEE Symposium on Information Visualization (InfoVis), Late Breaking Hot Topics*,. North Carolina, USA: IEEE, 1998. p. 9–12.
- ANSCOMBE, F. J. Graphs in statistical analysis. *The American Statistician*, v. 27, n. 1, p. 17–21, 1973.

ANTONIOU, G.; HARMELEN, F. V. *A semantic web primer*. [S.l.]: The MIT Press, 2004.

ARDITO, C. et al. Systematic inspection of information visualization systems. In: *BELIV '06: Proceedings of the 2006 AVI workshop on Beyond time and errors*. New York, NY, USA: ACM, 2006. p. 1–4. ISBN 1-59593-562-2.

ARONSON, A. Effective mapping of biomedical text to the UMLS Metathesaurus: the MetaMap program. In: AMERICAN MEDICAL INFORMATICS ASSOCIATION. *Proceedings of the AMIA Symposium*. [S.l.], 2001. p. 17–21.

ARRUARTE, A.; RUEDA, U.; ELORRIAGA, J. A. Organizing the learning resources related to the subject introduction to artificial intelligence through concept maps. In: *38th Annual Frontiers in Education Conference*. Saratoga Springs, New York, USA: IEEE Computer Society, 2008. p. F1C–17–F1C–22.

AVERBUKH, V. L. Towards the conceptions of visualization language and visualization metaphor. In: *Proceedings IEEE Symposia on Human-Centric Computing Languages and Environments*. Stresa, Italy: IEEE Computer Society, 2001. p. 390–391.

AVERBUKH, V. L. Visualization metaphors. *Program. Comput. Softw.*, v. 27, n. 5, p. 227–237, 2001.

AVERBUKH, V. L. et al. Interface and visualization metaphors. *Human-Computer Interaction. Interaction Platforms and Techniques*, p. 13–22, 2007.

AYDIN, G.; BALIM, A. G. Technologically - supported mind and concept maps prepared by students on the subjects of the unit "systems in our body". *Procedia - Social and Behavioral Sciences*, v. 1, n. 1, p. 2838–2842, 2009.

BAKER, J.; JONES, D.; BURKMAN, J. Using visual representations of data to enhance sensemaking in data exploration tasks. *Journal of the Association for Information Systems*, v. 10, n. 7, p. 2, 2009.

BARRAT, S.; TABBONE, S. Modeling, classifying and annotating weakly annotated images using bayesian network. *Journal of Visual Communication and Image Representation*, v. 21, n. 4, p. 355 – 363, 2010. ISSN 1047-3203.

- BARRETT, M. et al. Building bridges between local and global knowledge: new ways of working at the world health organisation. *Knowledge Management for Development Journal*, v. 1, n. 2, 2005.
- BATTISTA, G. D. et al. *Graph Drawing: Algorithms for the Visualization of Graphs*. New Jersey, USA: Prentice Hall, 1998.
- BECHHOFFER, S. *Ontoweb report: Ontology language standardisation efforts*. [S.l.], 2002.
- BEDERSON, B. B. et al. Zooming web browser. In: MARTIN, F.; PAUL, J.; HARRICK, M. V. (Ed.). *San Jose, CA, USA. Multimedia Computing and Networking*: SPIE, 1996. v. 2667, p. 260–271.
- BEEL, J.; GIPP, B. Enhancing search applications by utilizing mind maps. In: *Proceedings of the 21st ACM conference on Hypertext and hypermedia*. Toronto, Ontario, Canada: ACM, 2010. p. 303–304.
- BEMERS-LEE, T.; HENDLER, J.; LASSILA, O. The semantic web. *Scientific American*, 2001.
- BEPPLER, F. D. *Um Modelo para Recuperação e Busca de Informação Baseado em Ontologia e no Círculo Hermenêutico*. Tese (Doutorado), 2008.
- BERIO, G.; HARZALLAH, M. Knowledge management for competence management. *Journal of Universal Knowledge Management*, v. 0, n. 1, p. 21–28, 2005.
- BERIO, G.; HARZALLAH, M. Towards an integrating architecture for competence management. *Computers in Industry*, v. 58, p. 199–209, 2007.
- BERTIN, J. *Semiology of Graphics (trans. W. Berg)*. [S.l.: s.n.], 1983.
- BIOLCHINI, J. et al. Systematic review in software engineering. *System Engineering and Computer Science Department COPPE/UFRJ, Technical Report RT - ES 679/05*, v. 679, n. 05, 2005.
- BOARDMAN, R. Bubble trees the visualization of hierarchical information structures. In: *CHI '00 extended abstracts on Human factors in computing systems*. The Hague, The Netherlands: ACM, 2000. p. 315–316.

BODENREIDER, O. The unified medical language system (umls): integrating biomedical terminology. *Nucleic Acids Research*, Oxford Univ Press, v. 32, n. Database Issue, p. 267–270, 2004.

BODENREIDER, O.; ZHANG, S. Comparing the Representation of Anatomy in the FMA and SNOMED CT. In: AMERICAN MEDICAL INFORMATICS ASSOCIATION. *AMIA Annual Symposium Proceedings*. [S.l.], 2006. v. 2006, p. 46.

BONAMIGO, T.; SIQUEIRA, I. Screening for abdominal aortic aneurysms. *Rev. Hosp. Clin.*, v. 58, n. 2, p. 63–68, 2003.

BORGMAN, C. Why are online catalogs hard to use? Lessons learned from information-retrieval studies. *Journal of the American society for information science*, John Wiley & Sons, v. 37, n. 6, p. 387–400, 1986.

BOTTONI, P. et al. Storing and retrieving multimedia web notes. *Int. J. Comput. Sci. Eng.*, Inderscience Publishers, Inderscience Publishers, Geneva, SWITZERLAND, v. 2, n. 5/6, p. 341–358, 2006. ISSN 1742-7185.

BOWER, G. Cognitive psychology: an introduction. In: ESTES, W. (Ed.). *Handbook of learning and cognitive processes*. NJ: Hillsdale: Erlbaum, 1975. v. 1, p. 25–80.

BRESCIANI, S.; EPPLER, M. The benefits of synchronous collaborative information visualization: Evidence from an experimental evaluation. *IEEE Transactions on Visualization and Computer Graphics*, v. 15, n. 6, p. 1073–1080, nov.-dec. 2009. ISSN 1077-2626.

BRINKLEY, J. et al. A framework for using reference ontologies as a foundation for the semantic web. In: AMERICAN MEDICAL INFORMATICS ASSOCIATION. *AMIA Annual Symposium Proceedings*. [S.l.], 2006. v. 2006, p. 96–100.

BRULS, M.; HUIZING, K.; WIJK, J. V. Squarified treemaps. In: *Proceedings of the joint Eurographics and IEEE TCVG Symposium on Visualization*. Vienna: Springer, 2000. p. 33–42.

BUCHHEIM, C.; JÜNGER, M.; LEIPERT, S. Improving walker's algorithm to run in linear time. In: *Revised Papers from the 10th International Symposium on Graph Drawing*. Irvine, CA, USA: Springer-Verlag, 2002. p. 344–353.

BUDD, J. Mind maps as classroom exercises. *The Journal of Economic Education*, v. 35, n. 1, p. 35–46, 2004.

BUNGE, M. *Emergence and Convergence: Qualitative Novelty and the Unity of Knowledge (Toronto Studies in Philosophy)*. Canada: Univ of Toronto Press, 2003.

BURKHARD, R. et al. Beyond excel and powerpoint: knowledge maps for the transfer and creation of knowledge in organizations. In: *Ninth International Conference on Information Visualisation, 2005. Proceedings*. [S.l.: s.n.], 2005. p. 76 – 81. ISSN 1550-6037.

BURKHARD, R. A. Learning from architects: The difference between knowledge visualization and information visualization. In: *Eighth International Conference on Information Visualisation (IV'04)*. Los Alamitos, CA, USA: IEEE Computer Society, 2004. p. 519–524.

BURKHARD, R. A. *Knowledge Visualization – The Use of Complementary Visual Representations for the Transfer of Knowledge. A Model, a Framework, and Four New Approaches*. Tese (Doutorado) — Swiss Federal Institute of Technology Zurich, Switzerland, 2005.

BURKHARD, R. A. Strategy Visualization: A New Research Focus in Knowledge Visualization and a Case Study. In: *5th International Conference on Knowledge Management I-KNOW '05*. Graz, Austria: [s.n.], 2005.

BURKHARD, R. A. Towards a framework and a model for knowledge visualization: synergies between information and knowledge visualization. In: *Knowledge and Information Visualization*. Berlin/Heidelberg: Springer, 2005. v. 3426, p. 238–255.

BURKHARD, R. A.; MEIER, M. Tube map visualization: evaluation of a novel knowledge visualization application for the transfer of knowledge in long-term projects. *Journal of Universal Computer Science*, v. 11, n. 4, p. 473–494, 2005.

BUZAN, T.; BUZAN, B. *The mind map book*. 2. ed. London: BBC Books, 1995.

CADIZ, J. J.; GUPTA, A.; GRUDIN, J. Using web annotations for asynchronous collaboration around documents. In: *Proceedings of the 2000 ACM conference on Computer supported cooperative work*. Philadelphia, Pennsylvania, United States: ACM, 2000. p. 309–318.

CALVANESE, D.; LENZERINI, M. Ontology of integration and integration of ontologies. In: *Ontology of integration and integration of ontologies*. [S.l.: s.n.], 2001.

CAÑAS, A. et al. Concept maps: Integrating knowledge and information visualization. *Lecture Notes in Computer Science*, Springer, v. 3426, p. 205, 2005.

CARD, S. K.; MACKINLAY, J. D.; SHNEIDERMAN, B. *Readings in information visualization: using vision to think*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1999.

CARVALHO, E.; MARCOS, A. *Visualização de informação*. Portugal, 2009.

CASTRO P., L. A. Magnet mail: A visualization system for email information retrieval. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, v. 5531 LNCS, p. 213–222, 2009.

CERVO, A.; BERVIAN, P. *Metodologia científica: para uso dos estudantes universitários*. São Paulo: McGraw-Hill, 1983.

CHANDRASEKARAN, B. Design problem solving: a task analysis. *AI Mag.*, v. 11, n. 4, p. 59–71, 1990.

CHANDRASEKARAN, B.; JOHNSON, T.; BENJAMINS, V. Ontologies: What are they? why do we need them. *IEEE Intelligent Systems and their Applications*, v. 14, n. 1, p. 20–26, 1999.

CHANDRASEKARAN, B.; JOSEPHSON, J.; BENJAMINS, V. The ontology of tasks and methods. In: *Proceedings of the 11th Banff Knowledge Acquisition for Knowledge-Based System Workshop (KAW98)*. [S.l.: s.n.], 1998.

CHEN A.A., M. F. M.-C. C. B. K. H. Automated generation of individually customized visualizations of diagnosis-specific medical information using novel techniques of information extraction. In: . [S.l.: s.n.], 2005. v. 5748, p. 50–57.

CHEN, C. Top 10 unsolved information visualization problems. *Computer Graphics and Applications, IEEE*, v. 25, n. 4, p. 12–16, July-Aug. 2005. ISSN 0272-1716.

CHEN, C.; CZERWINSKI, M. P. Empirical evaluation of information visualizations: an introduction. *International Journal of Human-Computer Studies*, v. 53, n. 5, p. 631 – 635, 2000. ISSN 1071-5819.

CHEN, M. et al. Data, information, and knowledge in visualization. *Computer Graphics and Applications, IEEE*, v. 29, n. 1, p. 12–19, Jan.-Feb. 2009. ISSN 0272-1716.

CHEN, Y. The application of keystone concept map in scientific research and teaching. In: *First International Workshop on Education Technology and Computer Science*. Wuhan, China: IEEE, 2009. v. 2, p. 329–333.

CHITTARO, L.; COMBI, C. Representation of temporal intervals and relations: information visualization aspects and their evaluation. In: *Proceedings. Eighth International Symposium on Temporal Representation and Reasoning, TIME 2001*. [S.l.: s.n.], 2001. p. 13–20.

CHOI, N.; SONG, I.-Y.; HAN, H. A survey on ontology mapping. *SIGMOD Rec.*, v. 35, n. 3, p. 34–41, 2006.

CHRISTIAN, S.; PATRICK, E.; CHRISTOPHE, M. Information retrieval and visualization based on documents geospatial semantics. In: *International Conference on Information Technology: Research and Education, 2006. ITRE '06*. [S.l.: s.n.], 2006. p. 277–282.

CHRONAKI, C.; ZABULIS, X.; ORPHANOUDAKIS, S. Ic2 net medical image annotation service. *Informatics for Health and Social Care*, v. 22, n. 4, p. 337–347, 1997.

CHUNTING, Y.; YANG, L. Research on the application of concept map to software engineering teaching. In: *International Conference on Scalable Computing and Communications; Eighth International Conference on Embedded Computing*. Dalian, China: CPS, 2009. p. 612–615.

CLARKE, M.; OXMAN, A. *Cochrane Reviewers' Handbook 4.1.4 [updated October 2001]*. Oxford: The Cochrane Library, 2001.

COFFEY, J. W. et al. A concept map-based knowledge modeling approach to expert knowledge sharing. In: *Proceedings of IASTED International Conference on Information and Knowledge Sharing*. Virginia, USA: ACM, 2002.

- COFFEY, J. W.; HOFFMAN, R. R.; CAÑAS, A. J. Concept map-based knowledge modeling: perspectives from information and knowledge visualization. *Information Visualization*, v. 5, n. 3, p. 192–201, 2006.
- CONKLIN, J.; BEGEMAN, M. L. gibis: a hypertext tool for exploratory policy discussion. *ACM Trans. Inf. Syst.*, v. 6, n. 4, p. 303–331, 1988.
- COOK, M. P. Visual representations in science education: The influence of prior knowledge and cognitive load theory on instructional design principles. *Science Education*, v. 90, n. 6, p. 1073–1091, 2006.
- CORCHO, O. Ontology based document annotation: trends and open research problems. *International Journal of Metadata, Semantics and Ontologies*, v. 1, n. 1, p. 47–57, 2006.
- CORCHO, O.; PEREZ, A. A roadmap to ontology specification languages. In: *Knowledge Engineering and Knowledge Management. Methods, Models, and Tools*. [S.l.]: Springer Berlin / Heidelberg, 2000. p. 80–96.
- CRAFT, B.; CAIRNS, P. Beyond guidelines: what can we learn from the visual information seeking mantra? In: *Proceedings of Ninth International Conference on Information Visualisation*. [S.l.: s.n.], 2005. p. 110 – 118. ISSN 1550-6037.
- CRESWELL, J. *Research design: Qualitative, quantitative, and mixed methods approaches*. London, UK: Sage Publications, Inc, 2009.
- CRUBÉZY, M.; MUSEN, M. A. Ontologies in support of problem solving. In: *HANDBOOK ON ONTOLOGIES*. [S.l.]: Springer, 2003. p. 321–342.
- CUNNINGHAM, H. Information extraction - a user guide. *CoRR*, cmp-lg/9702006, 1997.
- CUNNINGHAM, H. Information extraction, automatic. *Encyclopedia of Language and Linguistics*, 2005.
- DACONTA, M. C.; SMITH, K. T.; OBRST, L. J. *The Semantic Web: A Guide to the Future of XML, Web Services, and Knowledge Management*. [S.l.]: John Wiley & Sons, Inc., 2003.

- DALEY, B. J.; TORRE, D. M. Concept maps in medical education: an analytical literature review. *Medical Education*, v. 44, n. 5, p. 440–448, 2010.
- DAVENPORT, T.; LONG, D. D.; BEERS, M. Successful knowledge management projects. *Sloan Management Review*, 1998.
- DAVIS, J. R.; HUTTENLOCHER, D. P. Shared annotation for cooperative learning. In: *The first international conference on Computer support for collaborative learning*. Indiana Univ., Bloomington, Indiana, United States: L. Erlbaum Associates Inc., 1995. p. 84–88.
- DING, G. et al. Automatic image annotations by mining web image data. In: *IEEE International Conference on Data Mining Workshops, 2009. ICDMW '09*. [S.l.: s.n.], 2009. p. 152–157.
- DING, G.; XU, N. Automatic semantic annotation of images based on web data. In: *2010 Sixth International Conference on Information Assurance and Security (IAS)*. [S.l.: s.n.], 2010. p. 317–322.
- DIX, A.; FINLAY, J.; ABOARD, G. *Human-computer interaction*. 3rd. ed. London: Prentice Hall, 2003.
- DOAN, A. et al. Learning to map between ontologies on the semantic web. In: *Proceedings of the 11th international conference on World Wide Web*. Honolulu, Hawaii, USA: ACM, 2002. p. 662–673.
- DOAN, A. et al. Ontology matching: A machine learning approach. In: STAAB, S.; STUDER, R. (Ed.). *Handbook on Ontologies*. [S.l.]: Springer Verlag, 2004. p. 385–516.
- DORK, M. et al. Visgets: Coordinated visualizations for web-based information exploration and discovery. *IEEE Transactions on Visualization and Computer Graphics*, v. 14, n. 6, p. 1205–1212, 2008.
- DRAPER, G. M.; LIVNAT, Y.; RIESENFELD, R. F. A survey of radial methods for information visualization. *IEEE Transactions on Visualization and Computer Graphics*, v. 15, n. 5, p. 759–776, 2009.
- EADES, P. D. Drawing free trees. *Bulletin of the Institute for Combinatorics and its Applications*, v. 5, p. 10–36, 1992.
- EBERT, D. S. et al. Two-handed volumetric document corpus management. *IEEE Computer Graphics and Applications*, v. 17, n. 4, p. 60–62, 1997.

EDWARD, C. Resultmaps: Visualization for search interfaces. *IEEE Transactions on Visualization and Computer Graphics*, v. 15, p. 1057–1064, 2009.

EHRIG, D. *Ontology Alignment: Bridging the Semantic Gap*. Tese (Doutorado), 2005.

EHRIG, M.; SURE, Y. Ontology mapping-an integrated approach. In: BUSSLER, C. et al. (Ed.). *The Semantic Web: Research and Applications*. [S.l.]: Springer Berlin / Heidelberg, 2004. v. 3053, p. 76–91.

ELAHI, N.; KARLSEN, R.; AKSELSEN, S. A context centric approach for semantic image annotation and retrieval. In: *Computation World: Future Computing, Service Computation, Cognitive, Adaptive, Content, Patterns, 2009. COMPUTATIONWORLD '09*. [S.l.: s.n.], 2009. p. 665 –668.

ELAHI, N.; KARLSEN, R.; YOUNAS, W. Image annotation by leveraging the social context. In: *Proceedings of the 5th International Conference on Ubiquitous Information Management and Communication*. New York, NY, USA: ACM, 2011. (ICUIMC '11), p. 40:1–40:8. ISBN 978-1-4503-0571-6.

ELAHI N., K. R. Y.-W. Semantic image annotation with social context. In: . [S.l.: s.n.], 2010.

ELVINS, T. T. A survey of algorithms for volume visualization. *SIGGRAPH Comput. Graph.*, v. 26, n. 3, p. 194–201, 1992.

EPPLER, M. Making knowledge visible through intranet knowledge maps: Concepts, elements, cases. In: *Proceedings of the 34th Annual Hawaii International Conference on System Sciences*. Maui, Hawaii, USA: IEEE Computer Society, 2001. v. 34, p. 4030.

EPPLER, M.; BURKHARD, R. A. Knowledge Visualization Towards a New Discipline and its Fields of Application. Università della Svizzera italiana, Faculty of Communication Sciences, Institute for Corporate Communication, 2004.

EPPLER, M.; BURKHARD, R. A. Knowledge Visualization. *Encyclopedia of Knowledge Management*, 2005.

EPPLER, M.; BURKHARD, R. A. Visual representations in knowledge management: framework and cases. *Journal of Knowledge Management*, v. 11, n. 4, p. 112–122, 2007.

- EVERS, H. Towards a Malaysian knowledge society. In: *Third International Malaysian Studies Conference (MSC3)*. Bangi: [s.n.], 2001.
- FABRIKANT, S. I.; MONTELLO, D. R.; MARK, D. M. The natural landscape metaphor in information visualization: The role of commonsense geomorphology. *Journal of the American Society for Information Science and Technology*, v. 61, n. 2, p. 253–270, 2010.
- FAHMI I., Z. J. E.-H. B. G. Swih system description: A case study in information retrieval, inference, and visualization in the semantic web. In: . [S.l.: s.n.], 2007. v. 4519 LNCS, p. 769–778.
- FARRAND, P.; HUSSAIN, F.; HENNESSY, E. The efficacy of the ‘mind map’ study technique. *Medical Education*, v. 36, n. 5, p. 426–431, 2002.
- FEKETE, J.-D. et al. The value of information visualization. In: *Information Visualization: Human-Centered Issues and Perspectives*. Berlin, Heidelberg: Springer-Verlag, 2008. p. 1–18.
- FELDMAN, R. et al. A comparative study of information extraction strategies. In: GELBUKH, A. F. (Ed.). *CICLing*. [S.l.]: Springer, 2002. (Lecture Notes in Computer Science, v. 2276), p. 349–359. ISBN 3-540-43219-1.
- FLATLA, D. R.; GUTWIN, C. Individual models of color differentiation to improve interpretability of information visualization. In: *Proceedings of the 28th international conference on Human factors in computing systems*. Atlanta, Georgia, USA: ACM, 2010. p. 2563–2572.
- FORSELL, C. A guide to scientific evaluation in information visualization. In: *14th International Conference Information Visualisation (IV)*. [S.l.: s.n.], 2010. p. 162–169. ISSN 1550-6037.
- FORSELL, C.; JOHANSSON, J. An heuristic set for evaluation in information visualization. In: *Proceedings of the International Conference on Advanced Visual Interfaces*. New York, NY, USA: ACM, 2010. (AVI '10), p. 199–206. ISBN 978-1-4503-0076-6.
- FOWLER, R. H.; FOWLER, W. A. L. Information visualization and retrieval using stereoscopic display of document and term relations. In: MARK, T. B.; SCOTT, S. F.; JOHN, O. M. (Ed.). *Stereoscopic*

Displays and Virtual Reality Systems V. San Jose, CA, USA: SPIE, 1998. v. 3295, p. 148–155.

FREITAS, C. et al. Evaluating usability of information visualization techniques. In: *Proc. 5th Symposium on Human Factors in Computer Systems (IHC) 2002, Brazilian Computer Society.* [S.l.: s.n.], 2002. p. 40–51.

FREITAS, J. *Um modelo de sistema de gestão do conhecimento para grupos de pesquisa e desenvolvimento.* Tese (Doutorado) — Programa de Pós-Graduação em Engenharia De Produção, Florianópolis, Brasil, 2003.

GAIZAUSKAS, R.; WILKS, Y. Information extraction: Beyond document retrieval. *Journal of Documentation*, MCB UP Ltd, v. 54, n. 1, p. 70–105, 1998.

GALLAGHER, R. *Computer Visualization: graphics techniques for scientific and engineering analysis.* Boca Raton, FL: CRC Press, 1995.

GARZOTTO, F.; PAOLINI, P.; SCHWABE, D. Hdm - a model-based approach to hypertext application design. *ACM Trans. Inf. Syst.*, ACM, New York, NY, USA, v. 11, p. 1–26, January 1993. ISSN 1046-8188.

GAUCH, H. *Scientific method in practice.* [S.l.]: Cambridge Univ Pr, 2003.

GERKEN, J.; JETTER, H.-C.; REITERER, H. Using concept maps to evaluate the usability of apis. In: *Proceedings of the 28th of the international conference extended abstracts on Human factors in computing systems.* Atlanta, Georgia, USA: ACM, 2010. p. 3937–3942.

GERSHON, N.; EICK, S. G.; CARD, S. Information visualization. *interactions*, v. 5, n. 2, p. 9–15, 1998.

GERSHON, N.; PAGE, W. What storytelling can do for information visualization. *Commun. ACM*, ACM, New York, NY, USA, v. 44, n. 8, p. 31–37, 2001. ISSN 0001-0782.

GIL, A. *Métodos e técnicas de pesquisa social.* São Paulo: Atlas, 1999.

GILCHRIST, A. Thesauri, taxonomies and ontologies-an etymological note. *Journal of Documentation*, v. 59, n. 1, p. 7–18, 2003.

GINSBURG, M. Annotate! a tool for collaborative information retrieval. In: *Proc. the 7th International Workshop on Enabling Technologies: Infrastructure for Collaborative Enterprises*. Palo Alto, California, USA: IEEE, 1998.

GREENGRASS, E. *Information retrieval: A survey*. 2000.

GRUBER, T. R. A translation approach to portable ontology specifications. *Knowledge Acquisition*, v. 5, p. 199–220, 2003.

GU, Q.; AHMAD, F.; SUMNER, T. Improving conceptual learning through customized knowledge visualization. In: *Third International Conference on Knowledge Discovery and Data Mining, 2010. WKDD '10*. [S.l.: s.n.], 2010. p. 407–410.

GUO, J.; LIAO, X. Cross-media image retrieval via latent semantic indexing and mixed bagging. In: *2009 WRI World Congress on Computer Science and Information Engineering*. [S.l.: s.n.], 2009. v. 4, p. 187–193.

HALASCHEK-WIENER, C. et al. A flexible approach for managing digital images on the semantic web. In: *Proceedings of the 5th International Workshop on Knowledge Markup and Semantic Annotation*. Galway, Ireland: CEUR-WS, 2005.

HALASCHEK-WIENER, C. et al. Photostuff—an image annotation tool for the semantic web. In: *Proceedings of the 4th International Semantic Web Conference*. Galway, Ireland: [s.n.], 2005.

HALASCHEK-WIENER, C. et al. Annotation and provenance tracking in semantic web photo libraries. In: *Provenance and Annotation of Data*. Heidelberg, Berlin: Springer, 2006, (Lecture Notes in Computer Science, v. 4145). p. 82–89.

HALASCHEK-WIENER, C. et al. Management of digital images on the semantic web. In: *Proceedings of the 4th International Semantic Web Conference*. Galway, Ireland: [s.n.], 2005.

HAN, J.; CERCONE, N. Ruleviz: a model for visualizing knowledge discovery process. In: *Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining*. New York, NY, USA: ACM, 2000. (KDD '00), p. 244–253. ISBN 1-58113-233-6.

- HAN, J.; CERCONE, N. Visualizing the process of knowledge discovery. *Journal of Electronic Imaging*, v. 9, p. 404, 2000.
- HANBURY, A. A survey of methods for image annotation. *Journal of Visual Languages & Computing*, v. 19, n. 5, p. 617–627, 2008.
- HARZALLAH, M.; BERIO, G. Competence modeling and management: A case study. In: *ICEIS (3)*. [S.l.: s.n.], 2004. p. 350–358.
- HARZALLAH, M.; BERIO, G.; VERNADAT, F. Analysis and modeling of individual competencies: Toward better management of human resources. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, v. 26, n. 1, p. 187–207, 2006.
- HASLHOFER, B.; KLAS, W. A survey of techniques for achieving metadata interoperability. *ACM Comput. Surv.*, v. 42, n. 2, p. 1–37, 2010.
- HAVERKAMP, D. S.; GAUCH, S. Intelligent information agents: Review and challenges for distributed information sources. *Journal of the American Society for Information Science*, v. 49, n. 4, p. 304–311, 1998.
- HEER, J.; CARD, S. K.; LANDAY, J. A. prefuse: a toolkit for interactive information visualization. In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. Portland, Oregon, USA: ACM, 2005. p. 421–430.
- HEIJST, G. van et al. Foundations for a methodology for medical kbs development. *Knowledge Acquisition*, v. 6, n. 4, p. 395–434, 1994.
- HEIMONEN, T.; JHAVERI, N. Visualizing query occurrence in search result lists. In: *Proceedings of the Ninth International Conference on Information Visualisation*. London, UK: IEEE Computer Society, 2005. p. 877–882.
- HEIMONEN, T.; SIIRTOLA, H. Visualizing query occurrence in mobile web search interfaces. In: *13th International Conference on Information Visualisation*. Barcelona, Spain: IEEE, 2009. p. 639–644.
- HERMAN, I.; MELANÇON, G.; MARSHALL, M. S. Graph visualization and navigation in information visualization: A survey. *IEEE Transactions on Visualization and Computer Graphics*, IEEE

- Educational Activities Department, Piscataway, NJ, USA, v. 6, n. 1, p. 24–43, 2000. ISSN 1077-2626.
- HÖFFLER, T. N.; LEUTNER, D. Instructional animation versus static pictures: A meta-analysis. *Learning and Instruction*, v. 17, n. 6, p. 722 – 738, 2007. ISSN 0959-4752.
- HOEBER, O.; YANG, X.; YAO, Y. Visiq: Supporting visual and interactive query refinement. *Web Intelligence and Agent Systems*, v. 5, n. 3, p. 311–329, 2007.
- HOLLINK, L. et al. Semantic annotation of image collections. In: HANDSCHUH, S. et al. (Ed.). *Proceedings of the Second International Conference on Knowledge Capture K-CAP 2003 on Knowledge Markup and Semantic Annotation*. Sanibel, Florida, USA: CEUR-WS, 2003. (Knowledge Capture 2003).
- HOU, J.; PAI, S. A spatial knowledge sharing platform. using the visualization approach. *International Journal of Production Research*, v. 47, n. 1, p. 25–50, 2009.
- HOW, R. D. et al. The commonkads organization model: Content, usage and computer support. *Expert Systems with Applications*, v. 11, n. 1, p. 29–40, 1996.
- HSU, W. et al. Spirs: A web-based image retrieval system for large biomedical databases. *International Journal of Medical Informatics*, v. 78, n. Supplement 1, p. S13–S24, 2009.
- HUBMANN-HAIDVOGEL, A.; SCHARL, A.; WEICHSELBRAUN, A. Multiple coordinated views for searching and navigating web content repositories. *Information Sciences*, v. 179, n. 12, p. 1813–1821, 2009.
- HYVÖNEN, E.; STYRMAN, A. Ontology-based image retrieval. In: HYVÖNEN, E.; KLEMETTINEN, M. (Ed.). *Proceedings of the XML Finland 2002 Conference*. Helsinki, Finland: Helsinki Institute for Information Technology (HIIT), 2002.
- INGWERSEN, P. *Information Retrieval Interaction*. [S.l.]: Taylor Graham, 1992.
- ISENBERG, P. et al. Grounded evaluation of information visualizations. In: *Proceedings of the 2008 conference on BEyond time and errors: novel evaluation methods for Information Visualization*. Florence, Italy: ACM, 2008. p. 1–8. ISBN 978-1-60558-016-6.

JACKSON, P. An exploratory survey of the structure and components of organizational memory. *Becoming Virtual: Knowledge Management and Transformation of the Distributed Organization*, p. 89–110, 2007.

JANSEN, B. J.; SPINK, A.; SARACEVIC, T. Real life, real users, and real needs: a study and analysis of user queries on the web. *Information Processing & Management*, v. 36, n. 2, p. 207–227, 2000.

JIANPING, Z.; DA, Z.; JIAHUA, Z. Knowledge visualization: An effective way of improving learning. In: *Second International Workshop on Education Technology and Computer Science (ETCS)*. Wuhan, China: IEEE, 2010. v. 1, p. 598–601.

JIANXIN, C. The using of mind map in concept design. In: *9th International Conference on Computer-Aided Industrial Design and Conceptual Design*. Kunming, China: IEEE, 2008. p. 1034–1037.

JIN Y., K. L. P.-B. Knowledge based image annotation refinement. *Journal of Signal Processing Systems*, v. 58, n. 3, p. 387–406, 2010. Cited By (since 1996) 1.

JOHN, D.; DIETER, F.; VAN, H. F. *Towards the Semantic Web: Ontology-driven Knowledge Management*. New York, NY, USA: John Wiley & Sons, Inc., 2003.

JOHNSON, B.; SHNEIDERMAN, B. Tree-maps: a space-filling approach to the visualization of hierarchical information structures. In: *Proceedings of the 2nd conference on Visualization '91*. San Diego, California: IEEE Computer Society Press, 1991. p. 284–291.

JR, C. E. K.; RUBIN, D. L. Automated semantic indexing of figure captions to improve radiology image retrieval. *Journal of the American Medical Informatics Association*, v. 16, n. 3, p. 380–386, 2009.

KAHAN, J. et al. Annotea: an open rdf infrastructure for shared web annotations. *Computer Networks*, v. 39, n. 5, p. 589–608, 2002.

KALFOGLOU, Y. Knowledge society arguments revisited in the semantic technologies era. *International Journal of Knowledge and Learning*, v. 3, n. 2, p. 225–244, 2007.

KALYUGA, S. Relative effectiveness of animated and static diagrams: An effect of learner prior knowledge. *Computers in Human Behavior*, v. 24, n. 3, p. 852–861, 2008.

KALYUGA, S.; CHANDLER, P.; SWELLER, J. Levels of expertise and instructional design. *Human factors*, v. 40, n. 1, 1998.

KELLER, T. et al. Information visualizations for knowledge acquisition: The impact of dimensionality and color coding. *Computers in Human Behavior*, v. 22, n. 1, p. 43–65, 2006.

KEPING, W.; XIAOJIE, W.; YIXIN, Z. A weighted feature support vector machines method for semantic image classification. In: *International Conference on Measuring Technology and Mechatronics Automation (ICMTMA 2010)*. [S.l.: s.n.], 2010. v. 1, p. 377–380.

KHAN, K. et al. Undertaking systematic reviews of research on effectiveness: Crd's guidance for those carrying out or commissioning reviews. 2001.

KIRSCHNER, P. A. Cognitive load theory: implications of cognitive load theory on the design of learning. *Learning and Instruction*, v. 12, n. 1, p. 1–10, 2002.

KIRYAKOV, A. et al. Semantic annotation, indexing, and retrieval. *Web Semantics: Science, Services and Agents on the World Wide Web*, v. 2, n. 1, p. 49–79, 2004.

KITCHENHAM, B. Procedures for performing systematic reviews. *Software Engineering Group, Department of Computer Science, Keele University and Empirical Software Engineering National ICT Australia Ltd, Keele University Technical Report, TR/SE-0401*, 2004.

KITCHENHAM, B. Guidelines for performing systematic literature reviews in software engineering. *Software Engineering Group, Department of Computer Science, Keele University and Empirical Software Engineering National ICT Australia Ltd, EBSE Technical Report, EBSE-2007-01*, 2007.

KNAUTZ, K.; SOUBUSTA, S.; STOCK, W. G. Tag clusters as information retrieval interfaces. In: *43rd Hawaii International Conference on System Sciences (HICSS)*. Poipu, Kauai, Hawaii: IEEE, 2010. p. 1–10.

KOSARA, R.; HAUSER, H.; GRESH, D. An interaction view on information visualization. In: *EUROGRAPHICS 2003 State-of-the-Art Reports*. [S.l.: s.n.], 2003. p. 123–137.

KROGH, G. V.; ROOS, J. *Organizational Epistemology*. New York, NY: St. Martin's Press, 1995.

KUMAR, P.; SUBRAMANIAN, N.; ZHANG, K. Evaluation of information visualization tools using the nfr approach. In: *Proceedings of the 10th international conference on Visual Information Systems: Web-Based Visual Information Search and Management*. Salerno, Italy: Springer-Verlag, 2008. p. 44–55.

LAKATOS, E. M.; MARCONI, M. d. A. *Fundamentos de metodologia científica*. São Paulo: Atlas, 1993.

LAM, A. Tacit knowledge, organizational learning and societal institutions: an integrated framework. *Organization Studies*, v. 21, n. 3, p. 487–513, 2000.

LEHMANN, S.; SCHWANECKE, U.; DÖRNER, R. Interactive visualization for opportunistic exploration of large document collections. *Information Systems*, v. 35, n. 2, p. 260–269, 2010.

LEHNER, F.; MAIER, R. K. How can organizational memory theories contribute to organizational memory systems? *Information Systems Frontiers*, Springer, v. 2, n. 3-4, p. 277–298, 2000.

LEIDE J.E., C. C. B.-J. L. A. Y. L. Task-based information retrieval: Structuring undergraduate history essays for better course evaluation using essay-type visualizations. *Journal of the American Society for Information Science and Technology*, v. 58, n. 9, p. 1227–1241, 2007.

LELAND, M. D. P.; FISH, R. S.; KRAUT, R. E. Collaborative document production using quilt. In: *Proceedings of the 1988 ACM conference on Computer-supported cooperative work*. Portland, Oregon, United States: ACM, 1988. p. 206–215.

LEGLER, R. Identifying the competencies of 'visual literacy' - a prerequisite for knowledge visualization. In: *Proceedings of Tenth International Conference on Information Visualization*. London, UK: IEEE, 2006. p. 232–236. ISSN 1550-6037.

LEUTNER, D.; LEOPOLD, C.; SUMFLETH, E. Cognitive load and science text comprehension: Effects of drawing and mentally imagining text content. *Computers in Human Behavior*, v. 25, n. 2, p. 284–289, 2009.

LÉVY, P. *Cyberculture*. MN, USA: Univ Of Minnesota Press, 2001. (Cyberculture traduzido por Robert Bononno).

LI, G.; LU, H. Visual knowledge recommendation service based on intelligent topic map. *Information Technology Journal*, v. 9, p. 1158–1164, 2010.

LI, Z. et al. Fusing semantic aspects for image annotation and retrieval. *Journal of Visual Communication and Image Representation*, 2010.

LIANG, J.; HUANG, M. L. Highlighting in information visualization: A survey. In: *14th International Conference on Information Visualisation (IV)*. [S.l.: s.n.], 2010. p. 79–85. ISSN 1550-6037.

LOBER, W. et al. Iml: An image markup language. In: *Proceedings, American Medical Informatics Association Fall Symposium*. [S.l.]: American Medical Informatics Association, 2001. p. 403–407.

LOPES, A.; FRACOLLI, L. Revisão sistemática de literatura e metassíntese qualitativa: considerações sobre sua aplicação na pesquisa em enfermagem. *Texto Contexto Enferm*, v. 17, n. 4, p. 771–8, 2008.

LOPES, M. et al. Pulmonary artery catheter complications: report on a case of a knot accident and literature review. *Rev Hosp Clin Fac Med Sao Paulo*, SciELO Brasil, v. 59, p. 77–85, 2004.

LU, J. et al. A new large-scale image automatic annotation system based on wordnet. In: *International Workshop on Education Technology and Training, 2008. and 2008 International Workshop on Geoscience and Remote Sensing. ETT and GRS 2008*. [S.l.: s.n.], 2008. v. 1, p. 758–762.

LUO, H. et al. Mining large-scale news video database via knowledge visualization. In: QIU, G. et al. (Ed.). *Advances in Visual Information Systems*. [S.l.]: Springer Berlin / Heidelberg, 2007, (Lecture Notes in Computer Science, v. 4781). p. 254–266.

LUO H., F. J. Y.-J. R. W. S. S. Analyzing large-scale news video databases to support knowledge visualization and intuitive retrieval. In: . [S.l.: s.n.], 2007. p. 107–114.

LYTRAS, M.; SICILIA, M. The knowledge society: a manifesto for knowledge and learning. *International Journal of Knowledge and Learning*, v. 1, n. 1, p. 1–11, 2005.

MAEDCHE, A. et al. Mafra - a mapping framework for distributed ontologies. In: *Proceedings of the 13th International Conference on Knowledge Engineering and Knowledge Management. Ontologies and the Semantic Web*. [S.l.]: Springer-Verlag, 2002. p. 235–250.

MARSHALL, C. C. Annotation: from paper books to the digital library. In: *Proceedings of the second ACM international conference on Digital libraries*. Philadelphia, Pennsylvania, United States: ACM, 1997. p. 131–140.

MARSHALL, C. C. Toward an ecology of hypertext annotation. In: *Proceedings of the ninth ACM conference on Hypertext and hypermedia : links, objects, time and space—structure in hypermedia systems: links, objects, time and space—structure in hypermedia systems*. Pittsburgh, Pennsylvania, United States: ACM, 1998. p. 40–49.

MARSHALL, C. C.; BRUSH, A. J. B. Exploring the relationship between personal and public annotations. In: *Proceedings of the 4th ACM/IEEE-CS joint conference on Digital libraries*. Tuscon, AZ, USA: ACM, 2004. p. 349–357.

MASHAYEKHI, V.; FEULNER, C.; RIEDL, J. Cais: collaborative asynchronous inspection of software. *SIGSOFT Softw. Eng. Notes*, v. 19, n. 5, p. 21–34, 1994.

MAYER, R.; MORENO, R. Nine ways to reduce cognitive load in multimedia learning. *Educational psychologist*, v. 38, n. 1, p. 43–52, 2003.

MAYER, R. E.; MORENO, R. Aids to computer-based multimedia learning. *Learning and Instruction*, v. 12, n. 1, p. 107–119, 2002.

MAYNARD, D.; BONTCHEVA, K.; CUNNINGHAM, H. Towards a semantic extraction of named entities. *Recent Advances in Natural Language Processing*, 2003.

MCGUFFIN, M. J.; ROBERT, J. M. Quantifying the space-efficiency of 2d graphical representations of trees. *Information Visualization*, v. 9, n. 2, p. 115–140, 2010.

MEDDES, J.; MCKENZIE, E. Improving visualization by capturing domain knowledge. In: ROBERT, F. E. et al. (Ed.). *Visual Data Exploration and Analysis VII*. San Jose, CA, USA: SPIE, 2000. v. 3960, p. 186–195.

MEENA, E.; KUMAR, A.; ROMARY, L. An extensible framework for efficient document management using rdf and owl. In: *Proceedings of the Workshop on NLP and XML (NLPXML-2004): RDF/RDFS and OWL in Language Technology*. Barcelona, Spain: Association for Computational Linguistics, 2004. p. 51–58.

MELGAR, A.; BEPPLER, F. D.; PACHECO, R. Knowledge retrieval in the anatomical domain. In: *Proceedings of the 1st ACM International Health Informatics Symposium*. Arlington, VA, USA: ACM, 2010.

MELGAR, A.; BEPPLER, F. D.; PACHECO, R. Um modelo para recuperação de informação na Área da saúde baseado em arquétipos visuais. In: *XII Congresso Brasileiro de Informática na Saúde - CBIS*. Porto de Galinhas, Pernambuco, Brasil: SBIS, 2010.

MELGAR, A.; BEPPLER, F. D.; PACHECO, R. A memória organizacional no contexto da engenharia do conhecimento. *DataGramaZero - Revista de Informação*, n. artigo aceito para publicação, 2011.

MELGAR, A. et al. A model for knowledge retrieval based on semantic images. In: *The Twenty-Third International Conference on Software Engineering and Knowledge Engineering (SEKE 2011)*. Miami, USA: [s.n.], 2011.

MELGAR, A.; BEPPLER, F. D.; PACHECO, R. C. Um modelo para a visualização do conhecimento baseado em arquétipos visuais. *Acta Scientiarum. Technology*, n. artigo aceito para publicação, 2011.

MELGAR, A.; PACHECO, R. Memória organizacional na Área da saúde: Uma abordagem a partir da engenharia do conhecimento e a visualização do conhecimento. In: *Memorias del Segundo Simposio Iberoamericano en Generación, Comunicación y Gerencia del Conocimiento en el contexto de la Novena Conferencia Iberoamericana en Sistemas, Cibernética e Informática*. Orlando, Florida, EE.UU.: International Institute of Informatics and Systemics, 2010.

MELGAR, A.; PACHECO, R. Recuperación de literatura biomédica usando imágenes semánticas. *Revista Electro Electrónica de la Pontificia Universidad Católica del Perú*, v. 3, p. 53–60, 2010.

MELGAR, A.; PACHECO, R. Software architecture for knowledge visualization from heterogeneous data sources in the anatomical

domain. In: *Proceedings of the IADIS international conference e-Society 2010*. Porto, Portugal: IADIS Press, 2010.

MELGAR, A.; PACHECO, R. Una arquitectura basada en ontologías para anotación semántica de imágenes en el dominio de la biomedicina. In: *Anales del CAIS 2010 - Congreso Argentino de Informática y Salud*. Ciudad Autónoma de Buenos Aires, Argentina: SADIO, 2010. p. 3148–3158.

MENDONÇA, E. A. et al. Accessing heterogeneous sources of evidence to answer clinical questions. *Journal of Biomedical Informatics*, v. 34, n. 2, p. 85–98, 2001.

MERRIAM, S. *Qualitative Research and Case Study Applications in Education*. San Francisco, USA: Jossey-Bass Publishers, 1998.

MICK, D. Consumer research and semiotics: Exploring the morphology of signs, symbols, and significance. *Journal of Consumer Research*, v. 13, n. 2, p. 196–213, 1986.

MILLER, G. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological review*, v. 63, n. 2, p. 81, 1956.

MILTON, N.; CLARKE, D.; SHADBOLT, N. Knowledge engineering and psychology: Towards a closer relationship. *International Journal of Human-Computer Studies*, v. 64, n. 12, p. 1214–1229, 2006.

MITRA, M.; CHAUDHURI, B. Information retrieval from documents: A survey. *Information Retrieval*, v. 2, n. 2-3, p. 141–163, 2000.

MOHD, W. M. B. W.; EMBONG, A.; ZAIN, J. M. A framework of dashboard system for higher education using graph-based visualization technique. In: ZAVORAL, F. et al. (Ed.). *Networked Digital Technologies*. [S.l.]: Springer Berlin Heidelberg, 2010, (Communications in Computer and Information Science, v. 87). p. 55–69. ISBN 978-3-642-14292-5.

MÖLLER, M.; SINTEK, M. A generic framework for semantic medical image retrieval. In: *Proceedings of the 1st International Workshop on Knowledge Acquisition from Multimedia Content KAMC'07*. Genoa, Italy: CEUR-WS.org, 2007.

MORRIS, C. Signs, language and behavior. Prentice-Hall, New York, USA, 1946.

MORSI, R.; IBRAHIM, W.; WILLIAMS, F. Concept maps: Development and validation of engineering curricula. In: *37th Annual Frontiers In Education Conference - Global Engineering: Knowledge Without Borders, Opportunities Without Passports*. Milwaukee, Wisconsin: IEEE, 2007. p. T3H-18-T3H-23.

MÜLLER, H. et al. A review of content-based image retrieval systems in medical applications-clinical benefits and future directions. *International Journal of Medical Informatics*, Elsevier, v. 73, n. 1, p. 1-23, 2004.

MULROW, C. *Systematic reviews: synthesis of best evidence for health care decisions*. Philadelphia: Amer College of Physicians, 1998.

NEVO, D.; WAND, Y. Organizational memory information systems: a transactive memory approach. *Decision Support Systems*, v. 39, n. 4, p. 549-562, 2005.

NEWSOM, D.; HAYNES, J. *Public relations writing: Form & style*. [S.l.]: Wadsworth Pub Co, 2007.

NGUYEN, T. N.; ZHANG, J. A novel visualization model for web search results. *IEEE Transactions on Visualization and Computer Graphics*, v. 12, n. 5, p. 981-988, 2006.

NIEDERHAUSER, D. et al. The influence of cognitive load on learning from hypertext. *Journal of Educational Computing Research*, v. 23, n. 3, p. 237-255, 2000.

NIELSEN, J.; MACK, R. *Usability Inspection Methods*. New York: John Wiley & Sons, 1994.

NIELSON, J. *Usability engineering*. Cambridge, MA: Academic Press, 1993.

NONAKA, I. A dynamic theory of organizational knowledge creation. *Organization science*, v. 5, p. 14-37, 1994.

NOVAK, J.; CAÑAS, A. The theory underlying concept maps and how to construct them. In: *Technical Report IHMC CmapTools 2006-01 Rev 01-2008*. Florida: Institute for Human and Machine Cognition, 2008.

NOVAK, J.; WURST, M. Collaborative knowledge visualization for cross-community learning. *Lecture notes in computer science*, Springer, v. 3426, p. 95, 2005.

NOWELL, L. T. et al. Visualizing search results: some alternatives to query-document similarity. In: *Proceedings of the 19th annual international ACM SIGIR conference on Research and development in information retrieval*. Zurich, Switzerland: ACM, 1996. p. 67–75.

NOY, N. et al. Pushing the envelope: challenges in a frame-based representation of human anatomy. *Data & Knowledge Engineering*, Elsevier, v. 48, n. 3, p. 335–359, 2004.

NOY, N. F.; DOAN, A.; HALEVY, A. Y. Semantic integration. *AI Mag.*, v. 26, n. 1, p. 7–9, 2005.

O'HARA, K.; SELLEN, A. A comparison of reading paper and on-line documents. In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. Atlanta, Georgia, United States: ACM, 1997. p. 335–342.

O'LEARY, D. Using ai in knowledge management: Knowledge bases and ontologies. *The Knowledge Management Yearbook 1999-2000*, p. 404, 1999.

ORGANIZATION, N. I. S. *Understanding Metadata*. Bethesda, MD, USA: NISO Press, 2004.

OSMAN, T. et al. An integrative semantic framework for image annotation and retrieval. In: *Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence*. [S.l.]: IEEE Computer Society, 2007. p. 366–373.

O'SULLIVAN, D.; WILSON, D.; BERTOLOTTO, M. Task-based annotation and retrieval for image information management. *Multimedia Tools and Applications*, p. 1–25, 2010.

PAEK, T.; DUMAIS, S.; LOGAN, R. Wavelens: a new view onto internet search results. In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. Vienna, Austria: ACM, 2004. p. 727–734.

PELTONEN, J. Visualization by linear projections as information retrieval. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, v. 5629 LNCS, p. 237–245, 2009.

PELTONEN, J. et al. An information retrieval perspective on visualization of gene expression data with ontological annotation. In:

2010 *IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)*. [S.l.: s.n.], 2010. p. 2178–2181. ISSN 1520-6149.

PéPIOT, G. et al. Uecml: Unified enterprise competence modelling language. *Comput. Ind.*, Elsevier Science Publishers B. V., Amsterdam, The Netherlands, The Netherlands, v. 58, n. 2, p. 130–142, 2007. ISSN 0166-3615.

PERRY, M.; AGARWAL, D.; MCPARLAND, C. Use scenarios for shared editing in scientific collaborations. In: *Proc the Fourth International Workshop on Collaborative Editing, CSCW2002 Conference*. New Orleans, Louisiana, United States: ACM, 2002.

PETTICREW, M.; ROBERTS, H. *Systematic reviews in the social sciences: A practical guide*. [S.l.]: Blackwell Publishing, 2005.

PILLAT, R. M.; VALIATI, E. R. A.; FREITAS, C. M. D. S. Experimental study on evaluation of multidimensional information visualization techniques. In: *Proceedings of the 2005 Latin American conference on Human-computer interaction*. New York, NY, USA: ACM, 2005. (CLIHIC '05), p. 20–30. ISBN 1-59593-224-0.

PLAISANT, C. The challenge of information visualization evaluation. In: *AVI '04: Proceedings of the working conference on Advanced visual interfaces*. New York, NY, USA: ACM, 2004. p. 109–116. ISBN 1-58113-867-9.

RAJAMANICKAM, V. *Infographics seminar handout*. BOMBAY, 2002.

<<http://www.albertocairo.com/infografia/noticias/2005/infographichandout.pdf>>
Acessado em 14 dez. 2010.

RAMACHANDRAN, S.; KASHI, R. An architecture for ink annotations on web documents. In: *Proceedings of the Seventh International Conference on Document Analysis and Recognition*. Edinburgh, UK: IEEE Computer Society, 2003. v. 1, p. 256.

RAMESH, B. Towards a meta-model for representing organizational memory. In: *The Thirtieth Hawaii International Conference on System Sciences*. [S.l.: s.n.], 1997. v. 2, p. 320–329 vol.2.

RANSDELL, J. Some leading ideas of peirce's semiotic. *Semiotica*, Walter de Gruyter, Berlin/New York, v. 19, n. 3-4, p. 157–178, 1977.

- RAUTER, A.; BENATO, K. Visualização da informação aplicada à estratégia competitiva de uma instituição educacional. In: *XXV Encontro Nac. de Eng. de Produção*. Porto Alegre, RS, Brasil: ENEGEP, 2005. p. 4618–4624.
- REINGOLD, E. M.; TILFORD, J. S. Tidier drawings of trees. *IEEE Transactions on Software Engineering*, SE-7, n. 2, p. 223–228, 1981.
- REN, L. Web image retrieval in web pages. In: *2010 2nd International Conference on Future Computer and Communication (ICFCC)*. [S.l.: s.n.], 2010. v. 1, p. V1–26 –V1–31.
- RIAZ, M.; MENDES, E.; TEMPERO, E. *Protocol for Systematic Literature Review of “Maintainability Prediction of Software Applications”*. [S.l.]: The University of Auckland New Zealand, 2008.
- ROBERTSON, G. G.; CARD, S. K.; MACKINLAY, J. D. Information visualization using 3d interactive animation. *Commun. ACM*, v. 36, n. 4, p. 57–71, 1993.
- RODRIGUES, J. F. *Design espacial-perceptivo: uma nova compreensão para representações visuais interativas*. Tese (Doutorado), 2007.
- ROSSE, C.; MEJINO, J. A reference ontology for biomedical informatics: the Foundational Model of Anatomy. *Journal of Biomedical Informatics*, Elsevier, v. 36, n. 6, p. 478–500, 2003.
- ROSSE, C.; MEJINO, J. The foundational model of anatomy ontology. In: *Anatomy Ontologies for Bioinformatics*. London, UK: Springer, 2008, (Computational Biology, v. 6). p. 59–117.
- ROSSE, C. et al. Motivation and Organizational Principles for Anatomical Knowledge Representation. *Journal of the American Medical Informatics Association*, v. 5, p. 17–40, 1998.
- ROSSE, C.; SHAPIRO, L.; BRINKLEY, J. The Digital Anatomist Foundational Model: Principles for Defining and Structuring its Concept Domain. In: *American Medical Informatics Association Fall Symposium*. [S.l.: s.n.], 1998. p. 820–824.
- RUBIN, D. et al. Medical imaging on the semantic web: Annotation and image markup. In: *AAAI Spring Symposium - Technical Report SS-08-05, pp. 93-98*. California, USA: AAAI Press, 2008.

RUBIN, D. et al. ipad: Semantic annotation and markup of radiological images. In: *AMIA Annu Symp Proc*. USA: American Medical Informatics Association, 2008. v. 2008, p. 626–630.

RUBIN, D. L. et al. Annotation and image markup: Accessing and interoperating with the semantic content in medical imaging. *Intelligent Systems, IEEE*, v. 24, n. 1, p. 57–65, 2009.

RYDBERG-COX J.A., V. L.-R. S. H. D. Approaching the problem of multi-lingual information retrieval and visualization in greek and latin and old norse texts. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, v. 3232, p. 168–178, 2004.

SANNOMIYA, T. et al. A framework for sharing personal annotations on web resources using xml. *Aust. Comput. Sci. Commun.*, v. 23, n. 6, p. 40–48, 2001.

SANTOS, J. H. V. *Considerações Acerca Dos Métodos Dedutivo E Indutivo*. Brasil: [s.n.], 2008.
<<http://pt.scribd.com/doc/10195328/Consideracoes-acerca-dos-metodos-dedutivo-e-indutivo>>. Acessado em 19 mai. 2011.

SAUSSURE, F. d. *Cours de Linguistique Générale*. Illinois, USA: Open Court Publishing Company, 1916. (Course in General Linguistics traduzido por Roy Harrys).

SCHREIBER, A. et al. Knowledge organisation in medical kbs construction. *Proc. of Medical Informatics Europe (MIE93)*, p. 394–405, 1993.

SCHREIBER, A. T. et al. Ontology-based photo annotation. *Intelligent Systems, IEEE*, v. 16, n. 3, p. 66–74, 2001.

SCHREIBER, G. et al. *Knowledge engineering and management: the CommonKADS methodology*. Cambridge Massachussets: The MIT Press, 1999.

SCHREIBER, G. et al. A mini-experiment in semantic annotation. In: HORROCKS, I.; HENDLER, J. (Ed.). *The Semantic Web - ISWC 2002*. Sardinia, Italia: Springer Berlin / Heidelberg, 2002. p. 404–408.

SCHWAMBORN, A. et al. Cognitive load and instructionally supported learning with provided and learner-generated visualizations. *Computers in Human Behavior*, v. 27, n. 1, p. 89–93, 2011.

SEBEOK, T. *Contributions to the Doctrine of Signs*. Indiana, USA: Indiana University Press, 1976.

SEIFERT I., E. K. H.-H. S. S. K. M.-R. N. N. G. Dilia - a digital library assistant: A new approach to information discovery through information extraction and visualization. In: . [S.l.: s.n.], 2009. p. 180–185.

SEKI, K.; MOSTAFA, J. Literature-based discovery by an enhanced information retrieval model. In: CORRUBLE, V.; TAKEDA, M.; SUZUKI, E. (Ed.). *Discovery Science*. Sendai, Japan: Springer, 2007. (Lecture Notes in Computer Science, v. 4755), p. 185–196. ISBN 978-3-540-75487-9.

SHNEIDERMAN, B. *Designing the user interface: strategies for effective human-computer interaction*. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 1986. ISBN 0-201-16505-8.

SHNEIDERMAN, B. Tree visualization with tree-maps: 2-d space-filling approach. *ACM Trans. Graph.*, v. 11, n. 1, p. 92–99, 1992.

SHNEIDERMAN, B. The eyes have it: a task by data type taxonomy for information visualizations. In: *Proceedings of the 1996 IEEE Symposium on Visual Languages*. Boulder, CO, USA: IEEE, Los Alamitos, CA, United States, 1996. p. 336–343.

SHNEIDERMAN, B.; PLAISANT, C. Strategies for evaluating information visualization tools: multi-dimensional in-depth long-term case studies. In: *BELIV '06: Proceedings of the 2006 AVI workshop on BEyond time and errors*. New York, NY, USA: ACM, 2006. p. 1–7. ISBN 1-59593-562-2.

SHVAIKO, P.; EUZENAT, J. *Ten Challenges for Ontology Matching*. [S.l.]: Springer-Verlag, 2008. 1164-1182 p.

SILVA, E. L. d.; MENEZES, E. M. *Metodologia da Pesquisa e Elaboração de Dissertação*. [S.l.]: UFSC, 2005.

SONDHI, P. et al. Question Processing and Clustering in INDOC: A Biomedical Question Answering System. *EURASIP Journal on Bioinformatics and Systems Biology*, Hindawi Publishing Corp. New York, NY, United States, v. 2007, n. 3, 2007.

SONG, G. Client-side visualization of internet forums for information retrieval. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, v. 5613 LNCS, n. PART 4, p. 606–613, 2009.

SONNTAG, D.; MÖLLER, M. A multimodal dialogue mashup for medical image semantics. In: *Proceeding of the 14th international conference on Intelligent user interfaces*. Hong Kong, China: ACM, 2010. p. 381–384.

SPEEL, P. et al. Conceptual modelling for knowledge based systems. *Encyclopedia of library and information science*, p. 16–42, 2002.

SPOERRI, A. Infocrystal: a visual tool for information retrieval & management. In: *Proceedings of the second international conference on Information and knowledge management*. Washington, D.C., United States: ACM, 1993. p. 11–20.

SPRING, M. et al. Embodying social capital facilitators in a collaborative authoring system. In: *Proc. Conference on Information Systems*. [S.l.: s.n.], 1997.

SRIVASTAVA, D.; VELEGRAKIS, Y. Intensional associations between data and metadata. In: *Proceedings of the 2007 ACM SIGMOD international conference on Management of data*. Beijing, China: ACM, 2007. p. 401–412.

STAAB, S. et al. Knowledge processes and ontologies. *Intelligent Systems, IEEE*, v. 16, n. 1, p. 26–34, 2001.

STASKO, J. et al. An evaluation of space-filling information visualizations for depicting hierarchical structures. *International Journal of Human-Computer Studies*, v. 53, n. 5, p. 663 – 694, 2000. ISSN 1071-5819.

STASKO, J.; ZHANG, E. Focus+context display and navigation techniques for enhancing radial, space-filling hierarchy visualizations. In: *IEEE Symposium on Information Visualization 2000*. Salt Lake City, Utah, USAC: IEEE Computer Society, 2000. p. 57–65.

STEFFEN, S.; ER, M.; SIEGFRIED, H. An annotation framework for the semantic web. In: ISJIZAKI, S. (Ed.). *Proceedings of the First Workshop on Multimedia Annotation*. Tokyo, Japan: [s.n.], 2001.

STEWART, T. A. *The Wealth of Knowledge: Intellectual Capital and the Twenty-first Century Organization*. [S.l.]: Doubleday, 2001.

STRANGE, A. Flickr document reveals origin of twitter. *Epicenter Blog*, 2007.
 <http://www.wired.com/epicenter/2007/04/flickr_document/>. Acesso em 31 jul. 2010.

STUDER, R.; BENJAMINS, V.; FENSEL, D. Knowledge engineering: Principles and methods. *Data & Knowledge Engineering*, Elsevier, v. 25, n. 1-2, p. 161–197, 1998.

STUDER, R. et al. Situation and perspective of knowledge engineering. In: CUENA, J. et al. (Ed.). *Knowledge Engineering and Agent Technology*. Amsterdam, The Netherlands: IOS Press, 2000. p. 237–232.

SU, G.; FENG, W.; MA, Y. A pairwise distances based visualization system for information retrieval. In: *2010 International Conference on Computational Intelligence and Software Engineering (CiSE)*. [S.l.: s.n.], 2010. p. 1–4.

SU, J. H. et al. Ontology-based semantic web image retrieval by utilizing textual and visual annotations. In: *Proceedings of the 2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology*. [S.l.]: IEEE Computer Society, 2009. v. 03, p. 425–428.

SUNASSE, N. N.; SEWRY, D. A. An investigation of knowledge management implementation strategies. In: *Proceedings of the 2003 annual research conference of the South African institute of computer scientists and information technologists on Enablement through technology*. [S.l.]: South African Institute for Computer Scientists and Information Technologists, 2003. p. 24–36.

SURE, Y.; STUDER, R. *On-to-knowledge methodology (OTKM)*. [S.l.], 2002.

SWELLER, J.; MERRIENBOER, J. van; PAAS, F. Cognitive architecture and instructional design. *Educational Psychology Review*, v. 10, n. 3, p. 251–296, 1998.

TARASSOV, V.; SANDKUHL, K.; HENOCH, B. Using ontologies for representation of individual and enterprise competence models. *International Conference on Research, Innovation and Vision for the Future*, p. 206–213, Feb. 12–16, 2006.

- TENG, C. T.; CHI, H. L. On visualization of cocitation networks. In: *Proceedings of the 11th International Conference Information Visualization*. Zurich, Switzerland: IEEE Computer Society, 2007. p. 470–475.
- TEOH, S. T.; MA, K.-L. Rings: A technique for visualizing large hierarchies. In: *Revised Papers from the 10th International Symposium on Graph Drawing*. Irvine, CA, USA: Springer-Verlag, 2002. p. 268–275.
- TERGAN, S. Digital concept maps for managing knowledge and information. *Knowledge and information visualization*, p. 185–204, 2005.
- THEODOSIOU, T.; ANGELIS, L.; VAKALI, A. Non-linear correlation of content and metadata information extracted from biomedical article datasets. *Journal of Biomedical Informatics*, v. 41, n. 1, p. 202–216, 2008.
- TRANFIELD, D.; DENYER, D.; SMART, P. Towards a methodology for developing evidence-informed management knowledge by means of systematic review. *British Journal of Management*, v. 14, n. 3, p. 207–222, 2003.
- TREISMAN, A. Preattentive processing in vision. *Computer Vision, Graphics, and Image Processing*, v. 31, n. 2, p. 156–177, 1985.
- TREVIÑOS, A. Introdução à pesquisa em ciências sociais: a pesquisa qualitativa em educação. *São Paulo: Atlas*, 1987.
- TUFTE, E.; HOWARD, G. *The visual display of quantitative information*. [S.l.]: Graphics press Cheshire, CT, 2001.
- UREN, V. et al. Semantic annotation for knowledge management: Requirements and a survey of the state of the art. *Web Semantics: Science, Services and Agents on the World Wide Web*, v. 4, n. 1, p. 14–28, 2006.
- USCHOLD, M. et al. The enterprise ontology. *Knowl. Eng. Rev.*, v. 13, n. 1, p. 31–89, 1998.
- VALIATI, E. R. A.; FREITAS, C. M. D. S.; PIMENTA, M. S. Using multi-dimensional in-depth long-term case studies for information visualization evaluation. In: *Proceedings of the 2008 conference on BEyond time and errors: novel evaluation methods for Information*

Visualization. New York, NY, USA: ACM, 2008. (BELIV '08), p. 9:1–9:7. ISBN 978-1-60558-016-6.

VARELA, F.; THOMPSON, E.; ROSCH, E. *The embodied mind: Cognitive science and human experience*. Cambridge, MA: [s.n.], 1991.

VASCONCELOS, J. et al. A Group Memory System for Corporate Knowledge Management: An Ontological Approach. *Proceedings of 1st European Conference on Knowledge Management (ECKM'2000)*, p. 91–99, 2000.

VATHANOPHAS, V.; SPRING, M. The use of peripheral awareness tools in collaborative systems. In: *Proc. the 11th Workshop on Information Technologies and Systems (WITS'01)*. New Orleans, Louisiana: [s.n.], 2001.

VEERASAMY, A.; BELKIN, N. J. Evaluation of a tool for visualization of information retrieval results. In: *Proceedings of the 19th annual international ACM SIGIR conference on Research and development in information retrieval*. Zurich, Switzerland: ACM, 1996. p. 85–92.

VENNA J., K. S. A.-H. N. K. P. J. Information retrieval perspective to nonlinear dimensionality reduction for data visualization. *Journal of Machine Learning Research*, v. 11, p. 451–490, 2010. Cited By (since 1996) 10.

VENZIN, M.; KROGH, G. von; ROOS, J. Future research into knowledge management. In: KROGH, G. V.; ROOS, J.; KLEINE, D. (Ed.). *Knowing in firms: Understanding, managing and measuring knowledge*. [S.l.: s.n.], 1998. p. 26–66.

VERA, D.; CROSSAN, M. Organizational learning and knowledge management: Toward an integrative framework. In: EASTERBY-SMITH, M.; LYLES, M. (Ed.). *The Blackwell handbook of organizational learning and knowledge management*. Malden, MA, USA: Wiley-Blackwell, 2003. p. 122–141.

WANG, H.; LIU, S.; CHIA, L.-T. Does ontology help in image retrieval?: a comparison between keyword, text ontology and multi-modality ontology approaches. In: *Proceedings of the 14th annual ACM international conference on Multimedia*. Santa Barbara, CA, USA: ACM, 2006. p. 109–112.

WANG, T. D. et al. Aligning temporal data by sentinel events: discovering patterns in electronic health records. In: *Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*. Florence, Italy: ACM, 2008. p. 457–466.

WARE, C. *Information visualization: perception for design*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2000.

WENG, C.; GENNARI, J. H. Asynchronous collaborative writing through annotations. In: *Proceedings of the 2004 ACM conference on Computer supported cooperative work*. Chicago, Illinois, USA: ACM, 2004. p. 578–581.

WIELINGA, B. J.; SCHREIBER, A. T.; GREEF, P. *Synthesis report. ESPRIT Project P1089, Deliverable Y3*. [S.l.]: University of Amsterdam, 1989.

WILLIS, C. L.; MIERTSCHIN, S. L. Mind maps as active learning tools. *J. Comput. Small Coll.*, v. 21, n. 4, p. 266–272, 2006.

WINCKLER, M. A.; PALANQUE, P.; FREITAS, C. M. D. S. Tasks and scenario-based evaluation of information visualization techniques. In: *Proceedings of the 3rd annual conference on Task models and diagrams*. New York, NY, USA: ACM, 2004. (TAMODIA '04), p. 165–172. ISBN 1-59593-000-0.

WOLFE, J. L. Effects of annotations on student readers and writers. In: *Proceedings of the fifth ACM conference on Digital libraries*. San Antonio, Texas, United States: ACM, 2000. p. 19–26.

WOLFE, J. M.; TREISMAN, A.; HOROWITZ, T. S. What shall we do with the preattentive processing stage: Use it or lose it? *Journal of Vision*, v. 3, n. 9, p. 572, 2003.

WORREN, N.; MOORE, K.; ELLIOTT, R. When theories become tools: toward a framework for pragmatic validity. *Human Relations*, v. 55, n. 10, p. 1227–1250, 2002.

WU, A. et al. Civil: support geo-collaboration with information visualization. In: *GROUP '09: Proceedings of the ACM 2009 international conference on Supporting group work*. New York, NY, USA: ACM, 2009. p. 273–276. ISBN 978-1-60558-500-0.

XIAO-YUE, W. Visualization based on concept maps: An efficient way to knowledge sharing and knowledge discovery in e-science

environment. In: YAN, M. (Ed.). *Fourth International Conference on Fuzzy Systems and Knowledge Discovery, FSKD 2007*. Haikou, Hainan, China: IEEE Computer Society, 2009. v. 2, p. 144–147.

XU H.-T., Z. X.-D. X. Y. S. B.-L. Adaptive model for web image semantic automatic annotation. *Ruan Jian Xue Bao/Journal of Software*, v. 21, n. 9, p. 2183–2195, 2010.

YANG, J.; WARD, M.; RUNDENSTEINER, E. Interring: An interactive tool for visually navigating and manipulating hierarchical structures. In: *Proceedings of IEEE Symposium on Information Visualization (InfoVis)*. Boston, Massachusetts, USA: [s.n.], 2002. p. 77–84.

YEE, K. et al. Animated exploration of dynamic graphs with radial layout. In: *IEEE Symposium on Information Visualization 2001*. San Diego, California: IEEE, 2001. p. 43–50.

YI, J. S. et al. Understanding and characterizing insights: how do people gain insights using information visualization? In: *Proceedings of the 2008 conference on BEyond time and errors*. New York, NY, USA: ACM, 2008. p. 1–6. ISBN 978-1-60558-016-6.

YOSHINAGA, N.; NOBUHARA, H. Formal concept analysis based web pages classification/visualization and their application to information retrieval. In: *2010 International Symposium on Communications and Information Technologies (ISCIT)*. [S.l.: s.n.], 2010. p. 153–157.

ZACK, M. Managing codified knowledge. *Sloan Management Review*, v. 40, n. 4, p. 45–58, 1999.

ZELENÝ, M. *Human systems management: integrating knowledge, management and systems*. MA, USA: World Scientific, 2005.

ZHANG, F. The application of visualization technology on knowledge management. In: *International Conference on Intelligent Computation Technology and Automation (ICICTA)*. Changsha, Hunan, China: IEEE Computer Society, 2008. v. 2, p. 767–771.

ZHANG, F.; DENG, S. Studies on the visualization for web information retrieval. In: *Wireless Communications, Networking and Mobile Computing, 2008. WiCOM '08. 4th International Conference on*. [S.l.: s.n.], 2008. p. 1–4.

ZHANG J., M. J. T.-H. Information retrieval by semantic analysis and visualization of the concept space of d-lib® magazine. *D-Lib Magazine*, v. 8, n. 10, 2002.

ZHENG, L. A novel framework for web image annotation. *Journal of Computational Information Systems*, v. 6, n. 1, p. 287–295, 2010.

ZHOU, X. et al. Automatic image annotation by an iterative approach: incorporating keyword correlations and region matching. In: *Proceedings of the 6th ACM international conference on Image and video retrieval*. Amsterdam, The Netherlands: ACM, 2007. p. 25–32.

ZHU, B.; WATTS, S. A. Visualization of network concepts: The impact of working memory capacity differences. *INFORMATION SYSTEMS RESEARCH*, v. 21, n. 2, p. 327–344, 2010.

ZIEMKIEWICZ, C.; KOSARA, R. The shaping of information by visual metaphors. *IEEE Transactions on Visualization and Computer Graphics*, v. 14, n. 6, p. 1269–1276, 2008.

ZUALKERNAN, I. A.; ABUJAYYAB, M. A.; GHANAM, Y. A. An alignment equation for using mind maps to filter learning queries from google. In: *Sixth International Conference on Advanced Learning Technologies*. Kerkrade, The Netherlands: IEEE Computer Society, 2006. p. 153–155.

ZUK, T.; CARPENDALE, S. Theoretical analysis of uncertainty visualizations. In: *Proc. SPIE & IS&T Conf. Electronic Imaging, Vol. 6060: Visualization and Data Analysis 2006*. [S.l.: s.n.], 2006. p. 66–79.

ZUK, T. et al. Heuristics for information visualization evaluation. In: *BELIV '06: Proceedings of the 2006 AVI workshop on BEyond time and errors*. New York, NY, USA: ACM, 2006. p. 1–6. ISBN 1-59593-562-2.