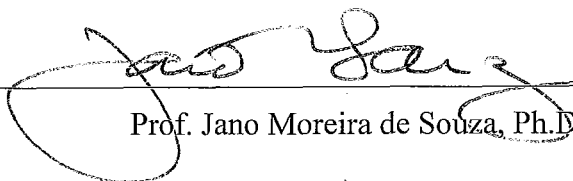


UTILIZAÇÃO DE TÉCNICAS DE VISUALIZAÇÃO PARA A RECOMENDAÇÃO
DE SUBSTITUTOS


Diogo Krejci de Souza

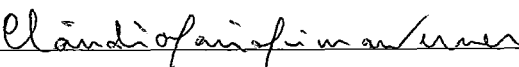
DISSERTAÇÃO SUBMETIDA AO CORPO DOCENTE DA COORDENAÇÃO DOS
PROGRAMAS DE PÓS-GRADUAÇÃO DE ENGENHARIA DA UNIVERSIDADE
FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS
NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE MESTRE EM CIÊNCIAS EM
ENGENHARIA DE SISTEMAS E COMPUTAÇÃO.

Aprovada por:


Prof. Jano Moreira de Souza, Ph.D.


Prof^a. Jonice de Oliveira Sampaio, D.Sc.


Prof^a. Claudia Lage Rebello da Motta, D.Sc.


Prof^a. Claudia Maria Lima Werner, D.Sc.

RIO DE JANEIRO, RJ – BRASIL

AGOSTO DE 2008

SOUZA, DIOGO KREJCI

Utilização de Técnicas de
Visualização para a Recomendação de
Substitutos [Rio de Janeiro] 2008

XV, 162., 29,7 cm (COPPE/UFRJ,
M.Sc., Engenharia de Sistemas e
Computação, 2008)

Dissertação – Universidade Federal
do Rio de Janeiro, COPPE

1. Gestão do Conhecimento
2. Conhecimento Científico
3. Visualização da Informação
4. Sistema de Recomendação

I. COPPE/UFRJ II. Título (série)

Agradecimentos

Gostaria de agradecer a todos os que me ajudaram direta ou indiretamente na realização desta dissertação.

A Deus, por todas as bênçãos que Ele me proporcionou até hoje.

Ao Professor Jano, meu orientador, por acreditar em mim e no meu trabalho.

A Jonice, que me orientou em todo o processo de preparação desse trabalho e que, mais que isso, sempre serviu como um exemplo de como devo agir para poder me tornar, um dia, um bom pesquisador e professor. Sem ela, esse trabalho certamente não seria o mesmo. A notícia de sua aprovação no concurso da UERJ foi a melhor notícia que recebi esse ano. Espero que mesmo longe nossa amizade permaneça.

Aos meus pais, Luiz e Rosângela, sempre presentes, apesar da distância. Não sei o que seria de mim sem vocês.

A minha irmã Laís e a Rose, minha segunda mãe, que muito me ajudaram sem mesmo perceber.

Ao meu tio Osvaldo Jambeiro que me deu todo suporte necessário no início da faculdade.

Ao meu primo Leo que me aturou durante 4 anos.

Ao meu primo André pelas partidas de sinuca para descontrair e as inúmeras caronas para Valença.

A minha tia Dulce pelas deliciosas refeições do fim de semana.

Aos meus amigos adquiridos durante o trabalho nos projetos da COPPETEC, Stainam, Ricardo, Sérgio, Tecla, Gustavo, Natália, Cadu, Luciana, Melissa, Monclar, Piter, Martino, Pedro, Severiano, Viviane, Miranda, Paulo Sérgio, Sômulo, Sávio e David, pela ajuda e momentos de descontração propiciados. Em especial, ao Leandro, muito mais que um amigo, quase um irmão, cuja presença no laboratório está fazendo uma grande falta.

Aos meus amigos adquiridos na graduação, Fred, Maurício, Vitor, Rafael, Renato e Luiz que facilitaram minha adaptação ao Rio de Janeiro.

Aos meus amigos de infância, que foram compreensivos nas minhas ausências devido à preparação deste trabalho.

As professoras Cláudia Motta e Cláudia Werner, por participarem da minha banca.

À Patrícia Leal, Carol e Vina, secretárias da linha de Banco de Dados do PESC, pela simpatia e atenção.

Aos funcionários da Secretaria do PESC, pela atenção e eficiência no atendimento às minhas muitas demandas operacionais.

Por fim, mas não menos importante, um agradecimento especial à minha namorada, Graziela, pelo carinho, companheirismo, paciência e compreensão durante os meus anos de mestrado.

Resumo da Dissertação apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências (M.Sc.)

UTILIZAÇÃO DE TÉCNICAS DE VISUALIZAÇÃO PARA A RECOMENDAÇÃO DE SUBSTITUTOS

Diogo Krejci de Souza

Agosto/2008

Orientadores: Jano Moreira de Souza
Jonice de Oliveira Sampaio

Programa: Engenharia de Sistemas e Computação

A busca por um pesquisador substituto adequado nas Instituições de Ensino e Pesquisa Brasileiras envolve a análise de uma grande quantidade de informação. Essa sobrecarga faz com que os responsáveis por essa busca gastem muito tempo processando grandes volumes de dados ou baseiem sua escolha na recomendação de terceiros. Esta dissertação propõe a abordagem BEE (Busca de Especialistas através de Exemplo) que, apoiada por técnicas de “Visualização da Informação”, indica substitutos a partir das competências e dos contextos de atuação do pesquisador a ser substituído. O uso de visualizações facilita o processo de substituição, pois proporciona uma representação clara e concisa dos valores dos parâmetros que o usuário considerou mais relevantes para comparação entre os substitutos candidatos. Esta abordagem é parte de um projeto de Gestão do Conhecimento Científico, o qual está sendo utilizado em instituições de ensino ou pesquisa no Brasil.

Abstract of Dissertation presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Master of Science (M.Sc.)

THE USE OF VISUALIZATION TECHNIQUES FOR SUBSTITUTE
RECOMMENDATION

Diogo Krejci de Souza

August/2008

Advisor: Jano Moreira de Souza.
Jonice de Oliveira Sampaio

Department: Systems and Computing Engineering

The search for an appropriated substitute researcher to be hired in the brazilian research institution involves the analysis of a large amount of information which means those responsible for it spend too much time processing huge volumes of data or base their choices on the recommendation of others. This dissertation proposes the BEE approach (BEE as the Portuguese acronym for a language meaning Search of Expert by Example). This tool indicates professionals to be hired through the support of Information Visualization techniques that assess competences and the working contexts of the researcher that is leaving a post. The visualization usage streamlines the replacement process as it provides a clear and concise representation of parameter values used during the assessment of applicants. This approach is part of the Scientific Knowledge Management Project, which is being used at some research institutes in Brazil.

Sumário

1.Introdução	1
1.1.Motivação	1
1.2.Objetivos da Dissertação	2
1.3.Contexto da Dissertação	2
1.4.Organização da Dissertação.....	3
2.Visualização da Informação	5
2.1.Introdução	5
2.2.Modelo de Referência para Visualização	7
2.2.1.Dado Bruto	8
2.2.2.Tabelas de Dados.....	8
2.2.3.Estruturas Visuais	11
2.2.4.Visões	16
2.3.Conclusão	18
3.Estado da Arte de Visualização da Informação.....	19
3.1.Motivação	19
3.2.Visualização de Espaços Complexos	20
3.2.1.Sistemas de Visualização de Classificação	21
3.2.1.1.Círculos Fractais	22
3.2.1.2.Mapas de Árvores (“TreeMaps”)	24
3.2.1.3.Árvore Hiperbólica (Hyperbolic Viewers).....	28
3.2.2.Sistemas de Visualização de Rotas.....	29
3.3.Sistemas de Visualização de Contextos.....	30
3.3.1.Sistemas de Exploração	31
3.3.1.1.Relevância Semântica.....	31

3.3.1.2.Consultas Dinâmicas	33
3.3.2.Sistemas de Mapa	35
3.3.2.1.Máquinas de Busca na Internet.....	35
3.3.2.2.Metáfora de Dados	37
3.3.2.3.Ambientes Virtuais	39
3.4.Visualização de Espaços Dinâmicos	39
3.4.1.Sistemas de Visualização de Fluxo	40
3.4.2.Sistemas de Evolução	41
3.5.Trabalhos Relacionados sobre Sistema de Recomendação	46
3.6.Conclusão	51
4.Busca de Especialistas através de Exemplo (BEE)	52
4.1.Introdução	52
4.2.Ambiente GCC	55
4.2.1.Módulo de Gestão de Competências	56
4.2.1.1.S-Miner: Mineração de Competências	56
4.2.1.2.Busca de Competências	57
4.2.1.3.Busca de Publicações	59
4.2.2.Módulo de Gestão de Conhecimento Pessoal	60
4.2.3.Módulo de Gestão de Projetos.....	60
4.2.4.Módulo de Comunidades.....	61
4.2.5.Módulo de Navegação e Visualização	62
4.2.6.Módulo de Recomendações.....	63
4.2.7.Módulo de Inferência.....	64
4.2.8.Módulo de Inteligência Competitiva	64
4.3.Descrição da Proposta	65
4.3.1.Ter a Informação na Base de Dados	67

4.3.2.Descobrir Contextos Críticos	69
4.3.3.Localizar Substitutos Aptos.....	79
4.3.3.1.Similaridade MBTI.....	91
4.3.3.2.Filtro de Aprendizado e Filtro de Interação com um Tema de Interesse	94
4.4.Comparação com Trabalhos Correlatos	98
4.5.Conclusão	103
5.Estudo de Caso	105
5.1.Definição	106
5.2.Planejamento	107
5.3.Execução.....	111
5.4.Análise dos Resultados.....	112
5.4.1.Perfil dos Participantes	113
5.4.2.Análise Comparativa da FBEE e das Ferramentas de Busca do GCC	113
5.4.2.1.Avaliação da Estrutura e Organização do Trabalho	114
5.4.2.2.Busca por Substitutos Através do GCC.....	114
5.4.3.Busca por Substitutos através da FBEE	120
5.4.4.Avaliação da FBEE	130
5.4.5.Conclusão Geral do Estudo Investigativo	133
6.Conclusões e Trabalhos Futuros.....	135
Referências Bibliográficas.....	139
Anexo I – Questionário de Avaliação do Estudo	148
Anexo II – Modelo de Dados Utilizado pela FBEE.....	160

Índice das Figuras

Figura 1 – Modelo de referência para visualização (CARD et al. 1999).	8
Figura 2 – Exemplo de dados brutos de um log de acesso fictício de um servidor web (CELMAR, 2006).	9
Figura 3 – Exemplos de marcas e algumas de suas propriedades. Traduzida de CARD et al. (1999).	12
Figura 4 – Visualização das despesas da.....	13
Figura 5 – Meta-estrutura do espaço de informação. Traduzido de JUDELMAN (2004)...	20
Figura 6 – Visualização exibida pelo sistema Grokker após ser efetuada a busca por times do futebol brasileiro (GROKKER, 2007).	23
Figura 7 – Visualização exibida pelo sistema Grokker após selecionar o círculo América Futebol Clube (GROKKER, 2007).	23
Figura 8 – Mapa de Árvore Cortar e Fatiar (SHNEIDERMAN, 1992).	25
Figura 9 – Aglomerada (WATTENBERG, 1999).	26
Figura 10 – Quadriculada (BRULS et al., 2000).	26
Figura 11 – Pivô por Tamanho e Pivô pelo Meio (BENJAMIN et al., 2002).	27
Figura 12 – Mapa de Árvore que representa a caixa de mensagens de um usuário (ADAM & MARC, 2006).	28
Figura 13 – Árvore CNPq, visão hiperbólica provida pelo GCC (OLIVEIRA et al., 2005).	28
Figura 14 – Árvore Bifocal. Adaptada de CAVA et al. (2002).	29
Figura 15 – Mapa produzido pela Opte Project em novembro de 2003 (OPTE PROJECT, 2007).	30
Figura 16 – Programa AutoFocus Personal (AUTOFOCOS PERSONAL, 2007).	32
Figura 17 – Programa Sinnzeug (SINNZEUG, 2007).	33
Figura 18 – Programa Film Finder (AHLBERG & SHNEIDERMAN, 1994).	34

Figura 19 – Exemplo do resultado retornado por uma busca realizada no programa Vivisimo (VIVISIMO, 2007).	36
Figura 20 – Exemplo do resultado retornado pela busca realizada no KARTOO (KARTOO, 2007).	36
Figura 21 – Exemplo da ferramenta WebSOM (WEBSOM, 2007).	38
Figura 22 – Exemplo de utilização do VXINSIGHT (KRISHNAN et al., 2007).	38
Figura 23 – Exemplo do arco de diagramas (WATTENBERG, 2002).	40
Figura 24 – Exemplo de visualização gerada pelo programa SeeNet3D (EICK et al., 1996).	41
Figura 25 – Imagem de uma visualização gerada pelo aplicativo ThemeRiver (HAVRE et al., 1999).	42
Figura 26 – Imagem de uma visualização gerada pelo Revisionist (REVISIONIST, 2007).	43
Figura 27 – Tabela periódica de métodos de visualização (LEGLER & EPPLER, 2007).	45
Figura 28 – “Expert landscape” para identificação de possíveis grupos (BECKS et al., 2004).	50
Figura 29 – Visão Geral da Arquitetura do GCC.	55
Figura 30 – Arquitetura do S-MINER (RODRIGUES et al., 2004).	57
Figura 31 – Fases da Busca da Competência (OLIVEIRA, 2007).	58
Figura 32 – Busca de Competências (OLIVEIRA, 2007).	59
Figura 33 – Árvore hiperbólica do GCC (OLIVEIRA et al., 2005).	63
Figura 34 – “Workflow” das etapas da abordagem BEE.	67
Figura 35 – Relevância das pessoas na organização.	70
Figura 36 – Relevância das pessoas no projeto.	79
Figura 37 – Visualização de Teia gerada na substituição de uma pessoa no contexto de projeto.	82

Figura 38 – Visualização de Área gerada na substituição de uma pessoa no contexto de projeto.....	83
Figura 39 – Visualização para Similaridade MBTI.....	93
Figura 40 – Gráfico utilizado na visualização do nível de conhecimento de uma competência no contexto projeto.....	95
Figura 41 – Informações do usuário no contexto organizacional.....	160
Figura 42 – Informações do usuário no contexto projeto.....	161
Figura 43 – Informações do usuário no contexto comunidade.....	162

Índice das Tabelas

Tabela 1 – Tabela de Dados organizando os dados brutos da Figura 2 (CELMAR, 2006).....	9
Tabela 2 – Dados derivados de uma tabela similar à	11
Tabela 3 – Tabela real de gastos de um residente de Goiânia (NASCIMENTO & FERREIRA, 2005).	13
Tabela 4 – Tabela de ordens de prioridade de MACKINLAY (1986).....	15
Tabela 5 – Efetividade relativa de diferentes propriedades retiniais. Traduzida de CARD et al. (1999).....	16
Tabela 6 – Competências do pesquisador P1 e seus respectivos graus.....	76
Tabela 7 – Competências do pesquisador P3 e seus respectivos graus.....	76
Tabela 8 – Competências do pesquisador P4 e seus respectivos graus.....	77
Tabela 9 – Competências do projeto Proj1 e sua utilização nos respectivos itens.....	77
Tabela 10 – Pesquisador P1 e sua participação no projeto Proj1.....	77
Tabela 11 – Pesquisador P3 e sua participação no projeto Proj1.....	77
Tabela 12 – Pesquisador P4 e sua participação no projeto Proj1.....	78
Tabela 13 – Cálculo do grau das competências do projeto e seus participantes.	90
Tabela 14 – Tipos de perfis de acordo com a tipologia MBTI MYERS (1995) e KEIRSEY & BATES (1984).....	92
Tabela 15 – Modelo de Domínio Intuitivo para Comparar Sistemas de Busca de Especialista. (YIMAM-SEID & KOBSA, 2003)	101
Tabela 16 – Informações detalhadas sobre os participantes do estudo.	113
Tabela 17 – Satisfação quanto a apresentação e estruturação do trabalho.	114
Tabela 18 – Itens utilizados na busca por substitutos no contexto de projetos.	115
Tabela 19 – Itens utilizados na busca por substitutos no contexto de comunidades.	115
Tabela 20 – Itens utilizados na busca por substitutos no contexto organizacional.	115

Tabela 21 – Funcionalidades utilizadas na busca por pesquisadores substitutos.....	117
Tabela 22 – Tempo gasto na busca por substitutos através do GCC.....	118
Tabela 23 – Substitutos recomendados.	119
Tabela 24 – Critérios utilizados na busca de por substitutos em projetos.....	121
Tabela 25 – Critérios utilizados na busca de por substitutos em comunidades.....	122
Tabela 26 – Critérios utilizados na busca de por substitutos na organização.....	123
Tabela 27 – Tempo gasto na busca por substitutos através da FBEE.	124
Tabela 28 – Substitutos recomendados através da FBEE.	125
Tabela 29 – Avaliação das visualizações providas pela FBEE.	127
Tabela 30 – Resultados referentes à comparação da busca realizada pela FBEE com a realizada através do GCC.	129
Tabela 31 – Tabela referente às respostas do questionário(Q5).....	130

Índice das Equações

Equação 1 – Relevância (cor) do pesquisador em um projeto.	72
Equação 2 – Área do retângulo do pesquisador em um projeto.	72
Equação 3 – Relevância (cor) do pesquisador em uma comunidade.	73
Equação 4 – Área do retângulo do pesquisador em uma comunidade.	74
Equação 5 – Relevância (cor) do pesquisador na organização.	75
Equação 6 – Área do retângulo do pesquisador na organização.	76
Equação 7 – Grau das competências do projeto.	85
Equação 8 – Grau das competências dos pesquisadores candidatos pertencentes ao projeto.	86
Equação 9 – Grau das competências da comunidade.	87
Equação 10 – Grau das competências dos pesquisadores candidatos pertencentes à comunidade.	88
Equação 11 – Grau das competências dos pesquisadores candidatos pertencentes à organização.	89

1. Introdução

Neste capítulo introdutório são apresentados a motivação, o problema a ser estudado e o objetivo deste trabalho. Neste capítulo também se encontra a organização do texto.

1.1. Motivação

A grande mobilidade no meio científico e acadêmico é preocupante. A falta de recursos e o baixo número de concursos fazem com que os pesquisadores permaneçam pouco tempo em seus projetos de pesquisa, pois são obrigados a ingressar no mercado ou trocar de instituição de ensino. Este problema não ataca somente recém-doutores ou pesquisadores que não conseguem entrar oficialmente para o corpo docente de uma instituição. Alunos de iniciação científica, mestrado e doutorado muitas vezes são atraídos por maiores salários do mercado, largando seus laboratórios e projetos de pesquisa. Desta forma, a instituição deve encontrar no seu próprio meio de pesquisa um pesquisador¹ apto a suprir essa perda.

No entanto, para chegar a esta pessoa, uma grande quantidade de informações tais como, competências, atividades acadêmicas, publicações, áreas de atuação etc., deve ser analisada. Essa sobrecarga de informação, somada a corriqueira apresentação textual da mesma, faz com que os responsáveis gastem muito tempo processando grandes volumes de dados ou baseiem sua escolha na recomendação de terceiros. Logo, na maioria das vezes, o pesquisador escolhido pode não ser o mais indicado. Devido a isso, torna-se necessário o uso de uma estratégia para organizar e representar essas informações, de modo que o esforço cognitivo necessário na sua análise seja minimizado.

A área de pesquisa denominada Visualização de Informação estuda maneiras de como o uso de computadores, na criação de representações visuais interativas a partir de dados abstratos, pode ampliar a cognição². Devido a essas representações visuais, o indivíduo passa a utilizar sua percepção visual na compreensão dos dados analisados, reduzindo assim o esforço cognitivo requerido no entendimento dos mesmos.

¹ Nesta dissertação, qualquer pessoa que faça parte do corpo docente ou discente de uma instituição e esteja realizando atividades de pesquisa será considerada um pesquisador.

² Cognição é o processo humano de uso e aquisição de conhecimento (MILLER, 1956).

A motivação para esta dissertação é que a utilização de técnicas de Visualização de Informação pode facilitar a análise de toda informação envolvida no processo de substituição de pesquisadores em seus contextos de atuação (projetos, comunidades e na própria instituição), de modo que a escolha do pesquisador apto a substituir um outro pesquisador seja a mais coerente possível e obtida em menor tempo.

1.2. Objetivos da Dissertação

O objetivo desta dissertação é definir uma abordagem para a recomendação de pesquisadores em instituições de ensino e pesquisa, de modo que a instituição escolha o melhor substituto possível. Essa abordagem orienta o usuário dentro do processo de substituição, pois revela a importância do pesquisador a ser substituído e possibilita a utilização de toda informação relevante à busca de um substituto.

É também objetivo dessa dissertação indicar algumas visualizações que auxiliem algumas etapas decisórias do processo de escolha de um substituto. Para isto, foi construída uma ferramenta para apoiar a abordagem citada. Essa ferramenta implementa uma busca de especialistas através exemplos, ou seja, o pesquisador a ser substituído é selecionado e, através das informações referentes a ele, todas as análises são realizadas. A ferramenta engloba técnicas de visualização que permitem a análise dos três espaços de informação propostos por JUDELMAN (2004): *complexidade* (“complexity”), *contexto* (“context”) e *dinâmica* (“dynamics”), os quais são explicados no capítulo 3 desta dissertação.

1.3. Contexto da Dissertação

Esta dissertação está inserida no contexto do cenário acadêmico e científico brasileiro. A solução proposta foi integrada ao GCC (Gestão de Conhecimento Científico), que é um ambiente “web” voltado ao público acadêmico, constituindo uma importante ferramenta para a Gestão do Conhecimento no ambiente científico (OLIVEIRA *et al.*, 2005).

O GCC possui todos os dados necessários à substituição de um pesquisador, mas não possui nenhuma ferramenta que a facilite. A busca por pesquisadores no GCC é feita de maneira textual, com isso muitas informações relevantes não são consideradas, como por

exemplo, a importância do pesquisador na instituição de ensino. Além disso, este ambiente não disponibiliza ferramentas que possibilite o agrupamento de mais que um atributo nos resultados decorrentes de uma busca por um substituto.

Esta abordagem, intitulada BEE (KREJCI *et al.*, 2008b), trouxe grandes contribuições para os usuários do GCC. Através da ferramenta que implementa essa abordagem, os usuários conseguem agrupar todas as informações sobre um pesquisador que são relevantes durante o processo de substituição. Isso permite que a substituição seja rápida e simples, através do uso de visualizações, colocando fim à busca textual que era realizada anteriormente. A ferramenta de BEE pode ser utilizada em outros contextos, como no meio empresarial, embora esse não seja o objetivo desta dissertação. Todas as informações e experimentos desta dissertação estão relacionados ao meio acadêmico. A busca por pesquisadores através da BEE será detalhada ao longo desta dissertação.

1.4. Organização da Dissertação

O restante deste trabalho está organizado da seguinte forma:

O Capítulo 2 apresenta fundamentos na área de Visualização de Informação, contendo modelos e conceitos utilizados na solução proposta pela dissertação e em sua implementação.

O Capítulo 3 apresenta o estado da arte de trabalhos relacionados à Visualização de Informação e a Sistemas de Recomendação de Especialista, uma análise das características desses trabalhos também é realizada.

O Capítulo 4 apresenta a solução proposta nesta dissertação, com base no uso de teorias de Visualização de Informação. Essa solução envolve a elaboração de uma abordagem para recomendação de substitutos e de uma ferramenta, a FBEE³, que viabiliza a aplicação desta abordagem. Neste capítulo também é apresentada uma análise comparativa da FBEE com os Sistemas de Recomendação de Especialistas apresentado no Capítulo 3.

³ FBEE é o protótipo funcional desenvolvido para avaliar a abordagem BEE, a partir deste momento sempre que nos referenciarmos a este protótipo o nome FBEE será utilizado.

O Capítulo 5 apresenta o projeto, execução e análise de um Estudo de Caso realizado para avaliar o comportamento da FBEE no apoio a recomendação de substitutos em um ambiente de Gestão do Conhecimento.

O Capítulo 6 apresenta as conclusões desta pesquisa, apresentando suas contribuições, limitações e trabalhos futuros.

O Anexo I apresenta os questionários aplicados na validação da abordagem proposta.

Por fim, o Anexo II apresenta o modelo de dados utilizado na implementação da BEE.

2. Visualização da Informação

Este capítulo apresenta conceitos fundamentais da área de Visualização de Informação, relacionados ao problema abordado por esta dissertação e à solução proposta para ele. A seção 2.1 apresenta uma introdução sobre o que é Visualização de Informação e o que ela pode fazer para auxiliar o processo humano de cognição e a seção 2.2 apresenta o modo como conjuntos de dados podem ser transformados em representações visuais, usando conceitos e técnicas da área.

2.1. Introdução

A proliferação da tecnologia de comunicação digital tem disponibilizado uma grande quantidade de informações de diversas origens e formatos. Usuários acessando essas grandes e diversificadas bases de dados ou realizando buscas na Internet obtêm facilmente um grande volume de informações, dentre as quais muitas podem ser irrelevantes para os objetivos da tarefa que está sendo realizada. Essa sobrecarga de informação faz com que as pessoas gastem muito tempo processando grandes volumes de dados na tentativa de extrair algum conhecimento significativo. Logo, torna-se necessário o uso de alguma estratégia para organizar e representar estas informações, de modo que os problemas gerados por sua sobrecarga sejam minimizados.

A capacidade cognitiva humana é um processo muito limitado, no referente ao tratamento de grandes volumes de informação (MILLER, 1956). Segundo NORMAN (1993), sem a ajuda externa, a memória, o pensamento e o raciocínio são todos restritos. Numa tentativa de auxiliar externamente a cognição, a humanidade vem desenvolvendo ao longo de sua história artefatos como ábacos, cartas de navegação e diagramas, como forma de registrar e expressar dados para serem posteriormente analisados por meio de nossos sentidos. Destes, a visão foi o sentido mais privilegiado, pois além de ser o primeiro componente do sistema sensorial, a visão é o sentido humano que apresenta maior capacidade de captação (CARD *et al.*, 1999). Através da visão, o ser humano pode reparar em detalhes e detectar padrões em dados representados pelos artefatos, trazendo à tona informações que seriam difíceis de analisar e entender caso fossem apresentadas de outras formas, por exemplo, na forma textual (WARE, 2004).

Desse modo, o uso da visão durante a análise de dados possibilita que a cognição expanda suas fronteiras para além da mente do indivíduo, estendendo-se para esses artefatos. Neste processo, denominado cognição externa (SCAIFE & ROGERS, 1996), o indivíduo representa através do artefato os dados que deseja analisar e usa sua percepção visual para compreender esses dados. Segundo CRAPO *et al.* (2000), um número de características visuais incluindo movimento, a cor, intensidade, tamanho, interseção, encerramento, orientação, direção da iluminação e distância do observador podem ser extraídas sem esforço consciente e num intervalo de 100-200 milissegundos. A cognição da mesma quantidade de informação, sem o apoio de alguma dessas características, pode levar centenas de milissegundos ou até mesmo minutos (BRAUTIGAM, 1996).

O uso do computador possibilitou um grande aprimoramento no processo de cognição externa. Sem o computador, uma pessoa criaria no papel um gráfico estático que seria elaborado conforme seu entendimento sobre uma determinada tarefa. Porém, a utilização do computador nesse processo viabiliza a adição de controles interativos ao gráfico, de forma que os observadores possam controlar os dados exibidos e organizá-los de modo a facilitar seu entendimento. Como as interações possíveis são definidas pelo projetista, cabe então a ele ter em mente os possíveis interesses dos usuários, para definir quais interações devem ser possibilitadas no gráfico (SPENCE, 2001).

A área de pesquisa denominada Visualização de Informação estuda maneiras de como o uso de computadores, na criação de artefatos visuais e interativos, pode facilitar o processo de cognição externa. Segundo CARD *et al.* (1999), a *Visualização de Informação* é o uso de uma representação visual, interativa e suportada por computador, de dados abstratos para ampliar a cognição. Com isso, o observador passa a utilizar a percepção para enxergar algum fenômeno nos dados, diminuindo assim seu esforço cognitivo.

Segundo CHEN & BÖRNER (2002), uma interface visual bem projetada:

1. Reduz o tempo de busca visual.
2. Viabiliza uma compreensão mais profunda do conjunto de dados.
3. Revela relacionamentos não notados de outra maneira.
4. Possibilita a visualização dos dados sob diferentes perspectivas.

Em outras palavras, a Visualização de Informação permite que o usuário consiga um maior número de “insights” sobre um conjunto de informações, alguns destes possíveis

“insights” são listados a seguir (SHNEIDERMAN & PLAISANT, 2005), (WEHREND & LEWIS, 1990), (WICKENS & HOLLANDS, 2000) e (ZHOU & FEINER, 1998):

1. Sumários: máximo, mínimo, média, porcentagem.
2. Busca: saber que item encontrar.
3. Padrões: distribuições, tendências, frequências, estruturas.
4. “Outliers”: exceções.
5. Relacionamentos: correlações, múltiplas interações.
6. “Tradeoff”: balanço, combinação de máximos ou mínimos.
7. “Clusters”: grupos, similaridades.
8. Caminhos: distâncias, conexões múltiplas.
9. Anomalias: erro nos dados.

Os dois primeiros “insights” são considerados simples, pois poderiam ser obtidos facilmente sem uso de algum tipo de representação visual. Entretanto, os demais “insights” são considerados complexos, pois sem o suporte de alguma representação visual, seria necessário um grande esforço cognitivo para descobri-los.

Por essas razões, a Visualização de Informação aparece como uma abordagem interessante para contornar os problemas inerentes da sobrecarga de informação, citados no início desta seção. Entretanto, para que uma representação visual dos dados seja efetiva para as pessoas, é fundamental que o mapeamento preserve o dado (CARD *et al.*, 1999). Esse processo de mapeamento será apresentado na próxima seção.

2.2. Modelo de Referência para Visualização

Um *modelo de referência para visualização* é o processo computacional que converte a informação numa visualização com a qual os usuários poderão interagir (CARD *et al.*, 1999), conforme mostrado na Figura 1. Nesse modelo, na primeira transformação trabalha-se o *dado bruto* e o armazena em *tabelas de dados*. O segundo passo consiste na escolha de uma *estrutura visual* (linear, mapa, tridimensional, temporal, árvore etc.) que melhor represente os dados armazenados. O último passo consiste nas *transformações visuais*, isto é, as possíveis interações que um usuário pode realizar sobre as estruturas visuais e com isso obter diferentes *visões* das mesmas. Os componentes desse modelo são descritos nas subseções a seguir.

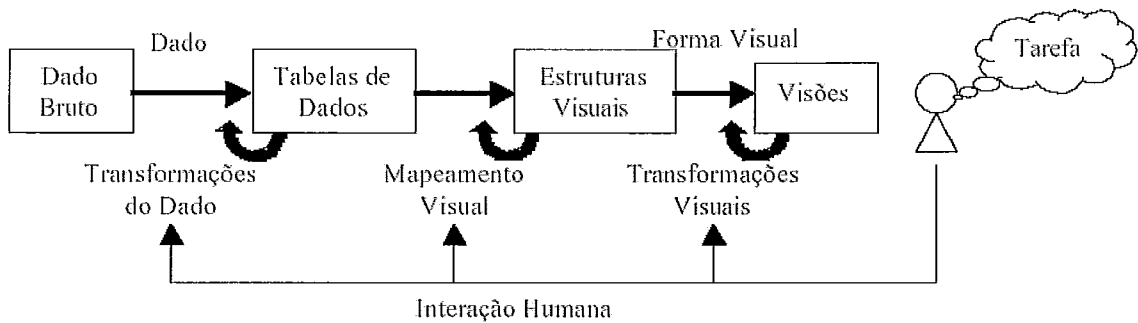


Figura 1 – Modelo de referência para visualização (CARD et al. 1999).

2.2.1. Dado Bruto

Na Figura 1, um indivíduo tem em mente a tarefa de estudar uma situação particular e obter informações relevantes sobre ela. Para tanto, primeiramente é necessário coletar dados sobre essa situação, para que a partir desses dados se consiga obter as informações necessárias. A coleta desses dados pode ser feita por sistemas computacionais ou manualmente pelo próprio indivíduo. Todo o dado coletado dessa forma recebe o nome de *dado bruto*.

Após a coleta dos dados brutos, torna-se necessário armazená-los em tabelas de dados. A organização dos dados brutos em tabelas é abordada a seguir.

2.2.2. Tabelas de Dados

O conjunto de dados brutos coletados pode ser heterogêneo, contendo datas, medidas, códigos, descrições, imagens, vídeos, endereços de “sites” da Internet, entre outros. Logo, diferentes entidades, atributos e relacionamentos podem estar representados nesse conjunto de dados.

Devido à heterogeneidade dos dados brutos, torna-se necessário organizá-los, de forma que alguma informação possa ser obtida pela consulta e comparação dos mesmos. Essa tarefa de tratamento de dados envolve organizar esses dados em estruturas (ou *tabelas de dados*), e é denominada *transformação de dados* (CARD et al., 1999). Quando os dados a serem tratados estão armazenados em um SGBD, essa transformação pode ser desnecessária, pois a organização do próprio banco de dados pode ser utilizada. Um

exemplo de dados brutos e seus respectivos mapeamentos em uma tabela de dados pode ser visto nas Figura 2 e na Tabela 1, respectivamente. Nesta tabela, cada coluna representa uma variável e cada linha representa um caso.

```

143.106.28.227 - silva [17/Feb/2003:16:53:41 -0300] "GET
/~silva/cursos/aplic/administracao/gerenciamento.php?
&cod_curso=22&acao=A HTTP/1.1" 200 3914
143.106.28.227 - silva [17/Feb/2003:16:53:43 -0300] "GET
/~silva/cursos/aplic/administracao/administracao.php?
&cod_curso=22 HTTP/1.1" 200 2403
143.106.28.227 - ferreira [17/Feb/2003:16:53:44 -0300] "GET
/~silva/cursos/aplic/administracao/gerenciamento.php?
&cod_curso=22&acao=N HTTP/1.1" 200 2940
143.106.28.227 - ferreira [17/Feb/2003:16:53:46 -0300] "GET
/~silva/cursos/aplic/administracao/gerenciamento.php?
&cod_curso=22&acao=C&ordem=nome HTTP/1.1" 200 4055
143.106.28.227 - silva [17/Feb/2003:16:53:48 -0300] "POST
/~silva/cursos/aplic/administracao/gerenciamento2.php HTTP/1.1" 200 2500
...

```

Figura 2 – Exemplo de dados brutos de um log de acesso fictício de um servidor web (CELMAR, 2006).

Tabela 1 – Tabela de Dados organizando os dados brutos da Figura 2 (CELMAR, 2006).

IP	Usuário	Tempo	Ação	Arquivo
143.106.28.227	silva	17/Feb/2003 16:53:41	GET	/~silva/cursos/aplic/administracao/gerenciamento.php
143.106.28.227	silva	17/Feb/2003 16:53:43	GET	/~silva/cursos/aplic/administracao/administracao.php
143.106.28.227	ferreira	17/Feb/2003 16:53:44	GET	/~silva/cursos/aplic/administracao/gerenciamento.php
143.106.28.227	ferreira	17/Feb/2003 16:53:46	GET	/~silva/cursos/aplic/administracao/gerenciamento.php
143.106.28.227	silva	17/Feb/2003 16:53:48	POST	/~silva/cursos/aplic/administracao/gerenciamento2.php

Cada atributo armazenado nas estruturas de dados pode ser classificado de acordo com sua semântica e com suas dependências funcionais em relação a outros atributos. Essas classificações são muito importantes, pois ajudam a escolher uma representação gráfica adequada para os dados, conforme será mostrado na 2.2.3.. Diferentes autores classificam de forma semelhante à semântica de atributos (ou variáveis) e de seus dados (SPENCE, 2001; CARD *et al.*, 1999; WARE, 2004), embora com nomenclaturas ligeiramente diferentes. Generalizando as diferenças, eles consideram a existência de três categorias básicas de atributos:

1. Nominiais (ou categóricos): conjunto de elementos sem ordem específica, como uma lista de títulos de filmes ou de nomes de pessoas.
2. Ordinais: conjunto de elementos distintos, mas que apresentam uma relação de ordem entre si, como classificações em faixas etárias, ou uma lista de datas históricas (neste caso, um atributo ordinal temporal).
3. Quantitativos (ou numéricos): conjunto de elementos que apresentam um escopo numérico, sendo possível aplicar operações aritméticas sobre eles. Essa categoria pode ser dividida em Intervalos (com valores discretos) e Razão (representando uma faixa contínua de valores). Por exemplo, distância entre cidades (atributo quantitativo espacial) ou quantidade de tempo que um usuário leva para ir de uma página “web” para outra em um “site” (atributo quantitativo temporal).

Cada atributo também pode ser classificado de acordo com sua dependência funcional relativa aos demais atributos. De acordo com essa característica, os atributos de uma tabela de dados podem ser divididos em atributos de entrada e atributos de saída, de tal forma que os primeiros determinam funcionalmente os últimos (de modo análogo a uma função matemática). Por exemplo, na Tabela 1 as variáveis IP, Usuário e Tempo são atributos de entrada, determinando funcionalmente os valores dos atributos de saída Ação e Arquivo.

Muitas vezes, o primeiro nível de organização de dados pode não ser suficiente para que o indivíduo obtenha informações relevantes sobre a situação estudada. Nesse caso, ordenação, cálculo de médias e classificação de características são algumas operações que podem ser aplicadas sobre as estruturas de dados obtidas. Essas novas transformações permitem classificar e agregar os dados que compartilham características semelhantes, possibilitando a sumarização e abstração dos mesmos (CARD et al., 1999). A Tabela 2 ilustra uma possível transformação aplicada sobre os dados da Tabela 1, mostrando valores derivados dos dados que ela apresenta.

Tabela 2 – Dados derivados de uma tabela similar à

Tabela 1 (CELMAR, 2006).

IP	Usuário	Quantidade de GETs	Quantidade de POSTs
143.106.28.227	silva	4	1
143.106.28.222	ferreira	9	0
143.106.28.234	ribeiro	1	0

Após o armazenamento do dado bruto em tabelas de dados, deve-se escolher uma estrutura visual que melhor represente os dados armazenados. Esse assunto é abordado a seguir.

2.2.3. Estruturas Visuais

Mesmo após diferentes operações de classificação e de agregação de dados, o conjunto de dados disponível ainda pode ser grande, tanto em quantidade quanto em número de dimensões, para se obter informação relevante. Isso acontece porque a representação de dados em formato textual exige que se faça um processamento visual controlado desses dados (CARD *et al.*, 1999), um processo que exige grande esforço cognitivo e que lida com uma pequena quantidade de dados por vez. Porém, as informações representadas graficamente tendem a serem processadas de maneira mais automática pela visão, em um processo mais superficial, paralelo (no sentido de obter várias informações simultaneamente), rápido e de capacidade elevada, diminuindo assim o esforço cognitivo necessário para se processar a informação.

Após organizar os dados que se deseja representar em uma tabela, é necessário determinar o que será usado para representá-los visualmente, definindo-se assim uma *estrutura visual*. Uma estrutura visual é um conjunto de elementos visuais que representam um conjunto de dados e pode ser decomposta em três partes:

1. Substrato espacial: é espaço para visualização, sendo normalmente representado por eixos, tais como os eixos X e Y do plano cartesiano.
2. Marcas: são objetos presentes no espaço do gráfico, como pontos, linhas, áreas, volumes (Card *et al.*, 1999), dígitos (Tufte, 2001) e ícones multidimensionais (Spence, 2001).

3. Propriedades gráficas das marcas: são os atributos visuais que caracterizam cada marca. Por exemplo, cores, formas, texturas, orientação, sua proximidade com relação a outras marcas etc. A Figura 3 mostra alguns exemplo de marcas e suas possíveis propriedades.

Dessa forma, para que uma estrutura visual represente os dados de uma tabela de dados, cada atributo dessa tabela pode ser associado a uma propriedade gráfica ou espacial, processo este chamado de *mapeamento visual*.

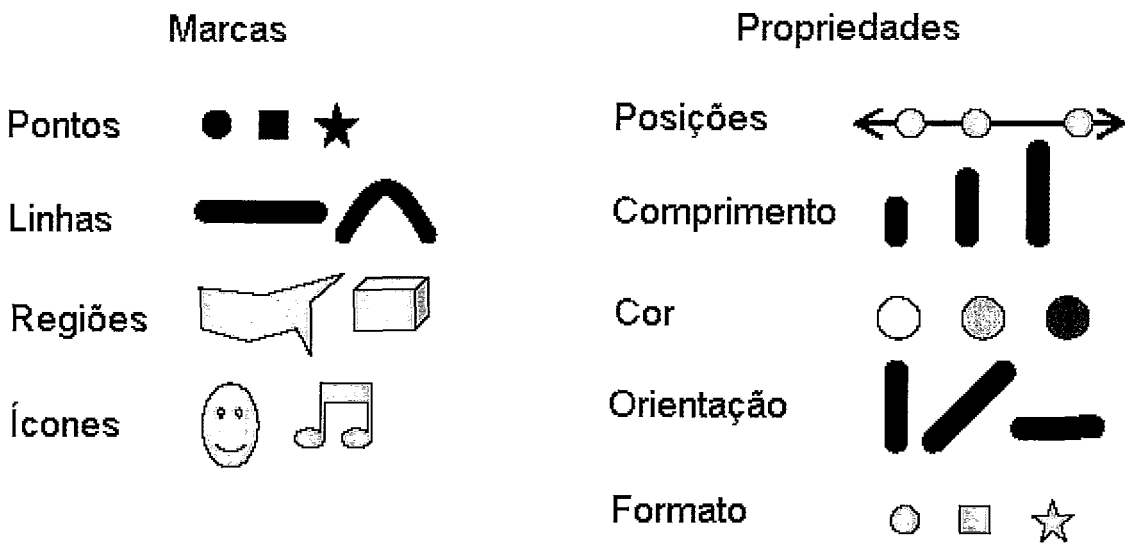


Figura 3 – Exemplos de marcas e algumas de suas propriedades. Traduzida de CARD et al. (1999).

Ao definir o mapeamento visual dos dados, o que se procura fazer é estabelecer como cada atributo será representado, ou seja, quais propriedades gráficas e espaciais podem ser utilizadas para sua representação. Por exemplo, a tabela de despesas, conforme mostrado na Tabela 3, foi mapeada no substrato espacial da Figura 4, contendo um eixo quantitativo (eixo Y) e um eixo ordinal (eixo X). Cada item de dados descrito por uma linha da tabela, foi associado a uma linha gráfica na Figura 4. Os atributos desses dados, neste caso o nome da despesa e seu valor em cada mês, foram associados à cor e à altura da linha em seus diversos pontos.

**Tabela 3 – Tabela real de gastos de um residente de Goiânia
(NASCIMENTO & FERREIRA, 2005).**

	Condomínio	Luz	Conta Telefone	Aluguel Telefone	Aluguel Apartamento	Faxina	Multicanal	Total
AGO	179,61	14,58	51,40	40,00	267,08	52,40		605,07
SET	183,81	23,50	38,35	40,00	267,08	52,40		605,14
OUT	201,21	30,24	149,00	40,00	267,08	52,40		739,93
NOV	219,73	35,94	143,95	40,00	232,08	52,40		724,10
DEZ	238,10	27,30	164,10	40,00	232,08	52,40		753,98
JAN	168,90	24,19	126,68	40,00	217,08	52,40		629,25
FEV	160,10	15,89	25,49	40,00	225,00	52,40		518,88
MAR	148,00	21,60	148,88	40,00	243,55	52,40		654,43
ABR	170,35	23,84	174,76	40,00	267,08	52,40		728,43
MAI	152,55	27,13	132,51	40,00	267,08	52,40		671,67
JUN	157,70	24,19	56,90	40,00	319,00	52,40		650,19
JUL	162,25	26,09	254,52	40,00	267,08	52,40		802,34
AGO	171,25	21,25	185,74	40,00	267,08	52,40	59,90	797,62
SET	155,85	29,55	114,42	40,00	267,08	52,40	59,90	719,20
OUT	148,90	28,68	171,74	40,00	265,00	52,40	59,90	766,62
NOV	150,35	15,38	98,16	40,00	265,00	52,40	57,90	679,19
DEZ	132,20	49,77	183,39	40,00	267,08	82,40	59,90	814,74
JAN	148,32	26,44	114,57	40,00	267,08	52,40	59,90	708,71

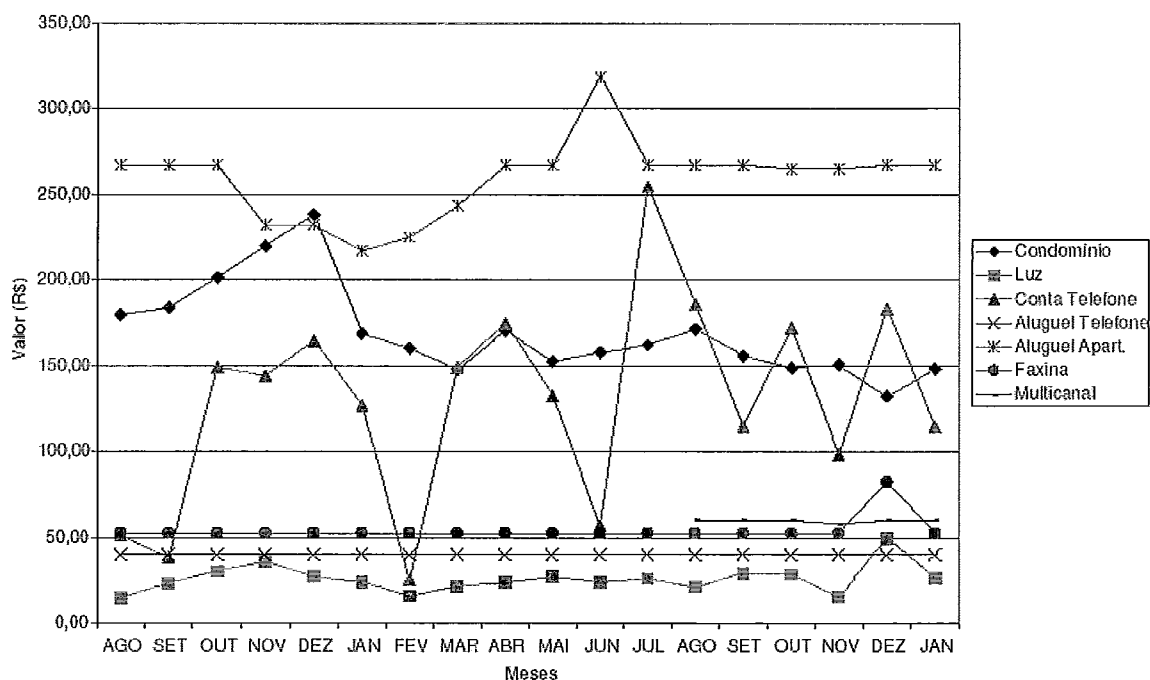


Figura 4 – Visualização das despesas da Tabela 3 por um gráfico de linha (NASCIMENTO & FERREIRA, 2005).

No exemplo acima, ficam evidentes alguns elementos com os quais se pode trabalhar em um mapeamento visual: espaço, marcas e propriedades gráficas dessas marcas (CARD *et al.*, 1999). Além destes, uma representação visual pode utilizar ainda o tempo

como uma forma de mapear outros atributos dos dados analisados. Para tanto, o indivíduo que efetua a análise pode utilizar controles interativos que filtram quais dados desses atributos devem ser considerados. Ao modificar os valores considerados e desconsiderados por esses filtros, esse indivíduo pode observar como são afetados os demais atributos já representados.

Entretanto, para se obter uma ampliação cognitiva através do *mapeamento visual*, é importante que este mapeamento seja feito da maneira apropriada. Um dos fatores que dificulta a definição de um mapeamento visual é a quantidade de propriedades gráficas disponíveis, como cores (e suas diferentes componentes, como tonalidade, saturação e brilho), formas (como diferentes polígonos ou ícones), tamanho, textura e orientação, dentre outros. Segundo MACKINLAY (1986), duas características devem ser consideradas ao se fazer um mapeamento visual:

1. *Expressividade* da estrutura visual: o mapeamento visual deve fazer com que a estrutura visual expresse todos os dados da tabela de dados, e somente eles.
2. *Efetividade* da estrutura visual: refere-se à capacidade de permitir rápida interpretação dos dados e fácil distinção entre eles, levando à menor quantidade possível de erros de interpretação.

Efetividade e expressividade são aspectos importantes, porque sem os mesmos, uma visualização pode não ser capaz de enfatizar padrões relevantes nos dados, não trazendo, assim, quaisquer informações novas além daquilo que já é trivialmente conhecido. Além disso, uma visualização também pode ser de difícil entendimento ou, até mesmo, sugerir padrões que na verdade não existem, o que pode levar a uma interpretação errônea dos dados (CARD *et al.*, 1999).

MACKINLAY (1986), com base nas três categorias de atributos (nominais, ordinais e quantitativos), propõe *ordens de prioridade* para se associar atributos de dados a atributos visuais, como uma forma de se produzir visualizações expressivas e efetivas. Essas ordens de prioridade vão do atributo visual mais perceptível até o menos perceptível. A idéia é associar os atributos mais importantes dos dados aos atributos visuais no topo da lista, seguindo a seqüência recomendada. A tabela de ordens de prioridades de MACKINLAY pode ser vista na Tabela 4.

Tabela 4 – Tabela de ordens de prioridade de MACKINLAY (1986).

Percepção	Dados Quantitativos	Dados Ordinais	Dados Nominais
Maior	Posição	Posição	Posição
	Comprimento	Densidade	Croma de Cor
	Ângulo	Saturação de cor	Textura
	Inclinação	Croma de cor	Conexão
	Área	Textura	Envolvimento
	Volume	Conexão	Densidade
	Densidade	Envolvimento	Saturação de cor
	Saturação de cor	Comprimento	Forma
	Croma de cor	Ângulo	Comprimento
	Textura	Inclinação	Ângulo
	Conexão	Área	Inclinação
	Envolvimento	Volume	Área
Menor	Forma	Forma	Volume

A tabela de efetividade de *propriedades retiniais* de CARD *et al.* (1999), que pode ser vista na Tabela 5, é outra abordagem que objetiva auxiliar a produção de visualizações expressivas e efetivas. Nessa tabela temos os seguintes significados: Q (dados quantitativos); O (dados ordinais); N (dados nominais). O símbolo ● representa que a propriedade é efetiva para mapear o tipo de dados a que aparece relacionado; ◐ indica que a propriedade é parcialmente Efetiva; e ○ indica que ela não é apropriada. Essa abordagem permite comparar a efetividade de propriedades gráficas com relação a cada categoria semântica de atributo (nominal, ordinal ou quantitativa), auxiliando, assim, a escolher quais categorias de atributos se deve representar por quais propriedades gráficas. Por outro lado, análises dos relacionamentos existentes entre eixos espaciais, marcas e suas propriedades gráficas também podem auxiliar na definição do mapeamento visual com relação à sua expressividade.

Tabela 5 – Efetividade relativa de diferentes propriedades retinais. Traduzida de CARD et al. (1999).

	Propriedades Espaciais	Q	O	N	Propriedades de Objetos	Q	O	N
Extensão de escala ¹	(Posição) Tamanho	● ●	● ●	● ●	Escala de cinza	○ ●	● ●	○ ○
Diferenciação de marcas	Orientação	○ ○	○ ○	● ●	Cor Textura Forma	○ ○ ○	○ ○ ○	● ● ●

O mapeamento de atributos em estruturas visuais, apresentado nesta seção, é considerado o núcleo do modelo de referência de Card *et al.* (1999), uma vez que nele reside o esforço de analisar as relações entre dados e transformá-las em propriedades gráficas efetivamente processadas pela visão humana.

Após a definição das estruturas visuais, inicia-se o último passo do *Modelo de Referência para Visualização* que são as *transformações visuais*, isto é, as possíveis interações que um usuário pode realizar sobre as estruturas visuais e com isso obter diferentes *visões* das mesmas. Este assunto é abordado a seguir.

2.2.4. Visões

As possíveis transformações visuais, chamadas de *transformações de visões*, que o usuário pode realizar nas estruturas das visualizações é outro fator importante. Essas transformações modificam as estruturas visuais interativamente, transformando apresentações estáticas em visualizações que estabelecem parâmetros gráficos para criação de *visões* das estruturas visuais. Com isso, o usuário pode modificar dinamicamente qual a fatia do conjunto de dados a ser exibida, reduzindo o excesso de dados e auxiliando seu processo de obtenção de informação. Além disso, o esforço de exploração dos dados é repassado em parte para o computador. Uma vez que os cálculos e o redesenho da imagem são realizados pela máquina, fica a cargo do usuário apenas a tarefa de observar o que acontece quando a visualização se modifica.

As técnicas desenvolvidas para a Visualização de Informação envolvem formas de representação do dado (2D, 3D, árvores), interação humana (consultas dinâmicas,

manipulação direta) e as tarefas. SHNEIDERMAN (1996) propôs uma taxonomia que prevê sete tipos de tarefas:

1. Visão geral: compreende uma visão geral da coleção inteira. Essa visão geral contém uma caixa de campo de visão móvel com a qual o usuário controla o conteúdo da visão detalhada. As interfaces de Visualização de Informação normalmente suportam estratégias de visão geral da coleção.
2. Afastamento e aproximação: a tarefa de afastamento e aproximação deve estar sempre presente nas visualizações. Sempre que o usuário tem interesse em alguma parte da coleção, ele precisará de ferramentas para controlar o foco de afastamento ou aproximação sobre a parte desejada.
3. Filtro: a tarefa de filtragem permite ao usuário remover da visualização os itens que não lhe interessam no momento, focando seu interesse na informação relevante. Essa tarefa geralmente é implementada através de consultas dinâmicas sobre os que fazem parte da coleção.
4. Detalhe sob demanda: uma vez que a coleção tenha sido reduzida a poucos itens, deve ser possível obter detalhes sobre esses itens. Geralmente usa-se o “mouse” para selecionar um item e com isso obter-se as informações dos atributos desse item numa janela.
5. Relação: permitir que o usuário perceba a relação entre os itens da visualização é outra tarefa importante em uma visualização. Porém, é rara que uma simples ação de um usuário produza a saída desejada. Por esse motivo, a criação de ações de interface para exibir relações é um desafio.
6. História: é importante manter a história das ações realizadas durante a exploração de informação, no caso dessa exploração ser feita em muitos passos e o usuário desejar refazer esses passos.
7. Extração: a extração de sub-coleções deve ser outra tarefa possível de se realizar. Uma vez que o usuário tenha obtido o item ou conjunto de itens que deseja, seria útil que ele pudesse extrair esse conjunto e salvá-lo num formato de arquivo, para posteriormente poder enviá-lo por correio eletrônico, imprimi-lo, transformá-lo em gráfico ou inseri-lo num pacote de cálculos estatísticos.

A partir do momento que o usuário passa a utilizar uma estrutura visual, dois fatores devem ser considerados: a *usabilidade* e as *funcionalidades* dessa estrutura. ROCHA & BARANAUSKAS (2003) definem usabilidade como a facilidade que um usuário final tem de aprender a usar o sistema, a eficiência que ele utiliza o sistema após aprendê-lo e o quão agradável é a utilização desse sistema. Mesmo a usabilidade sendo um fator crítico na definição de como será uma estrutura visual, ela ainda não recebe a atenção devida. Os estudos referentes à usabilidade ainda são muito lentos e se baseiam, em sua maioria, em metodologias que antecedem a área de Visualização de Informação (CHEN, 2005). Geralmente essas metodologias não são adequadas para tratar de detalhes críticos à área de Visualização de Informação, como a necessidade de verificar se os padrões apresentados aos usuários, através de uma estrutura visual, serão reconhecidos por eles corretamente. CHEN (2005) também destaca que as questões referentes à usabilidade ainda são tratadas de uma maneira “ad-hoc”, isto é, limitadas ao sistema que se tem em mãos.

2.3. Conclusão

Neste capítulo foi apresentada a área de Visualização de Informação e alguns de seus conceitos e técnicas, os quais podem auxiliar programadores a desenvolverem estruturas visuais interativas, e que sejam adequadas para ajudar usuários em seus processos de análise visual de dados e de tomada de decisão. O próximo capítulo apresenta o estado da arte de algumas ferramentas de visualização da informação e de recomendação de especialistas, visto que o objetivo deste trabalho é propor uma abordagem que aplica técnicas de Visualização da Informação na recomendação de substitutos.

3. Estado da Arte de Visualização da Informação

Neste capítulo é apresentado um estado da arte de sistemas de visualização da informação e de sistemas de recomendação de especialistas. O primeiro se faz necessário, visto que, a análise de sistemas de visualização da informação é a melhor forma de se obter idéias para um novo sistema de visualização. Enquanto o segundo está relacionado ao objetivo da dissertação, que é facilitar a busca por especialistas aptos a substituir uma outra pessoa.

A seção 3.1 apresenta a motivação para realização do estado da arte referente a Visualização da Informação. As seções 3.2, 3.3 e 3.4 apresentam os três espaços de informação propostos por JUDELMAN (2004) e que forma utilizados na categorização das visualizações encontradas. Por fim, a seção 3.5 apresenta um estado da arte sistemas de recomendação de especialistas, além da comparação dos sistemas considerados.

3.1. Motivação

A natureza da informação, as ações permitidas ao usuário e a aplicabilidade do sistema a ser produzido são determinantes na definição da visualização a ser produzida. Interfaces (bidimensional ou tridimensional) freqüentemente representam muitos atributos (tempo, tamanho, localização etc.) de uma informação em uma imagem, tornando-se necessário otimizar essas interfaces para certos usuários com certas necessidades. Logo, através da análise de muitos sistemas de visualização, pode-se obter um ponto inicial para novas idéias de como fazer essa otimização.

O estado da arte de *Visualização da Informação* apresentado neste capítulo se baseou na classificação proposta por JUDELMAN (2004), dado que a maioria dos sistemas de visualização foca uma das categorias mostradas pelo autor. Com base nas distinções da cartografia e na *Ultra Estrutura* proposta por LONG & DOROTHY (1995), JUDELMAN (2004) identificou três aspectos do espaço de informação que podem ser percebidos visualmente: caminhos complexos, conteúdo relacional e transformações dinâmicas. Respectivamente chamados: *complexidade* (“complexity”), *contexto* (“context”) e *dinâmica* (“dynamics”), como pode ser observado na Figura 5. Esses três termos descrevem as meta-estruturas mais expostas por sistemas de visualização e são detalhados nas seções a seguir.

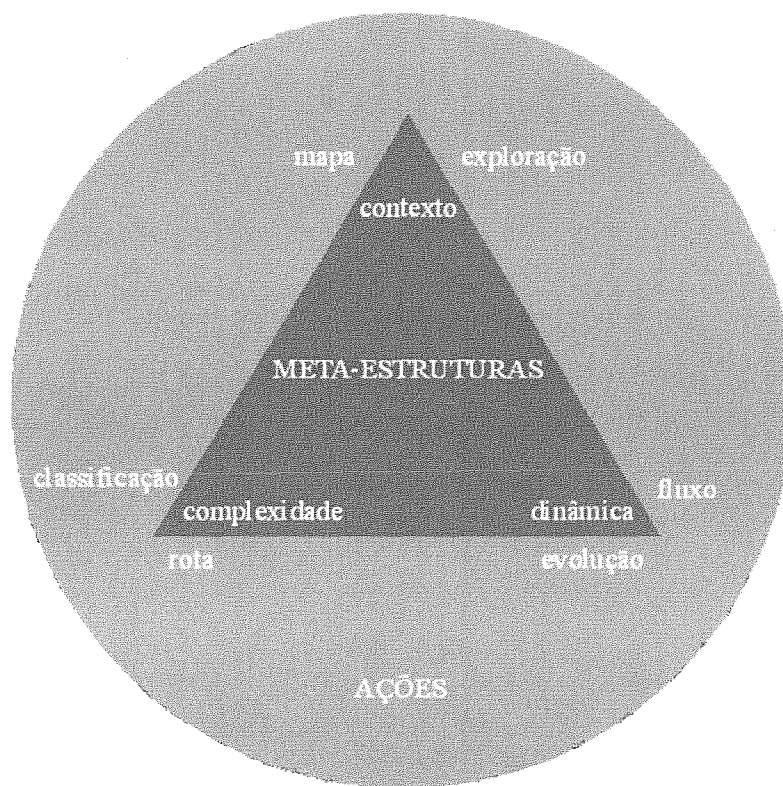


Figura 5 – Meta-estrutura do espaço de informação. Traduzido de JUDELMAN (2004).

3.2. Visualização de Espaços Complexos

Um dos focos da Visualização de Informação é examinar técnicas e projetos que facilitem a navegação e orientação do usuário dentro de uma estrutura complexa. *Espaços Complexos*, tais como árvores e hierarquias, mostram a topologia ou caminho da estrutura da informação e são geralmente retratados por imagens de redes com muitos nós (JUDELMAN, 2004). Portanto, o grande desafio é exibir uma grande quantidade de elementos e como eles estão ligados, de modo que o usuário tenha o seu entendimento sobre esse espaço de informação facilitado.

Nesta seção, os exemplos de sistemas de visualização estão divididos em duas sub-categorias:

1. Classificação (“classify”): sistemas de visualização utilizados na navegação através hierarquias.

2. *Rota* (“route”): sistemas de visualização utilizados para exibição da topologia da estrutura da informação.

3.2.1. Sistemas de Visualização de Classificação

Um exemplo comum de estrutura de informação complexa é a classificação hierárquica, isto é, uma categorização de elementos que se ramificam em sub-elementos. O objetivo dos sistemas de visualização de classificação é facilitar a navegação do usuário através de hierarquias complexas de informação. Usuários que trabalham com hierarquias podem ter diversas tarefas, logo a visualização dessas hierarquias pode ajudar a responder as seguintes perguntas (CARD *et al.*, 1999):

1. Qual o número total de nós da hierarquia?
2. Quantos nós se ramificam em sub-nós?
3. Qual a distância entre o nó mais profundo e o nó raiz?
4. Quais são os nomes de cada nó?
5. Quais são os atributos de cada nó?
6. Qual nível da hierarquia possui mais nós?

Ao se construir sistemas de classificação, um grande problema se faz presente: como permitir ao usuário navegar através das hierarquias sem que este perca a noção da estrutura geral da mesma, este problema é conhecido como *problema do buraco da fechadura* (“Keyhole Problem”).

SHNEIDERMAN & PLAISANT (2005) propõem um padrão de projeto para visualizações que objetiva solucionar o *problema do buraco da fechadura*. Primeiramente, deve-se fornecer ao usuário uma visão geral do espaço de informação, em detrimento do detalhamento das mesmas. A seguir, disponibilizar ao usuário mecanismos de interação, de modo que este possa fazer um “zoom” sobre a informação desejada ou retirar, da visualização, as informações desnecessárias no momento. Por último, prover uma visão detalhada da informação selecionada pelo usuário. “*Overview first, zoom and filter, then detail on demand*” (SHNEIDERMAN & PLAISANT, 2005).

A visão geral do espaço de informação pode ser obtida através do uso da técnica de *visão geral e detalhe* (“overview and detail”) ou de *foco e contexto* (“focus and context”). Ambas as técnicas provêm uma visão detalhada da estrutura da informação, sem que o

usuário perca a noção de todo o espaço de informação. A técnica de *visão geral e detalhe* utiliza uma janela para exibir a visão geral da informação e outra para exibir os detalhes da informação selecionada. Enquanto a técnica de *foco e contexto* procura integrar visão geral e detalhes numa única janela, um exemplo de aplicação dessa técnica pode ser visto na Figura 7.

A seguir são listadas algumas das vantagens de se fornecer, primeiramente, ao usuário uma visão geral do espaço de informação:

1. Facilita a formação de modelos mentais do espaço de informação.
2. Revela qual informação está presente ou não na visualização.
3. Exibe relacionamentos entre partes da informação, favorecendo um maior número de “insights”.
4. Permite acessar diretamente ou navegar por partes da informação, bastando para isso selecionar a informação desejada na visão geral.
5. Estimula a exploração.

Nas subseções a seguir são apresentados alguns exemplos de sistemas de classificação.

3.2.1.1. Círculos Fractais

Círculos fractais é uma técnica utilizada para exibir hierarquias de informação. A hierarquia dos círculos é composta por objetos dentro de objetos, onde o número de círculos é igual ao número de objetos que possuem sub-objetos. O uso de cores nessa técnica pode ser usado para diferenciar tipos e tamanhos de objetos.

O sistema de visualização GROKKER (GROKKER, 2007) usa a metáfora dos círculos fractais para exibir a hierarquia da informação juntamente com a técnica de *foco e contexto*. A primeira visualização exibida pelo sistema GROKKER mostra uma visão geral do espaço de informação hierárquico (*foco*), como na Figura 6. Uma visão detalhada dos sub-objetos contidos em cada círculo pode ser obtida selecionando-se em um desses círculos (contexto), vide Figura 7. Para desfazer a operação basta clicar na opção “zoom back”.

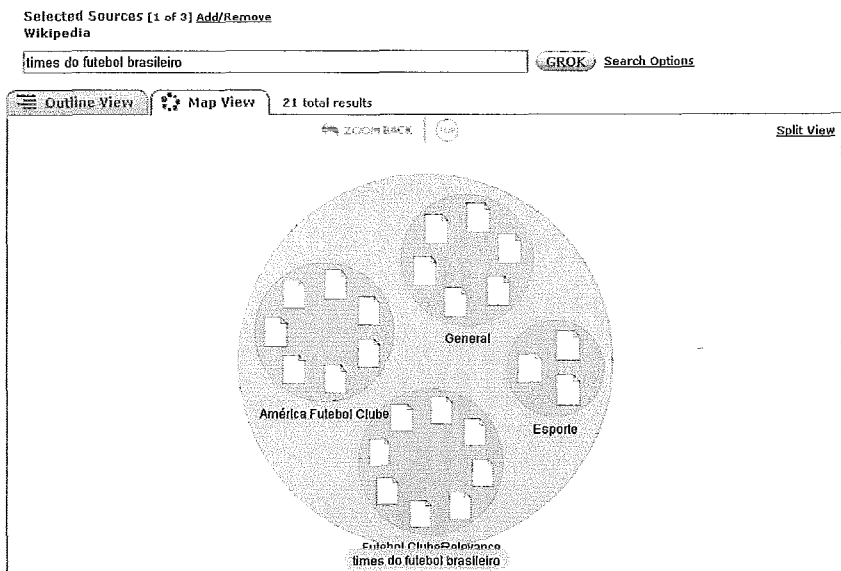


Figura 6 – Visualização exibida pelo sistema Grokker após ser efetuada a busca por times do futebol brasileiro (GROKKER, 2007).

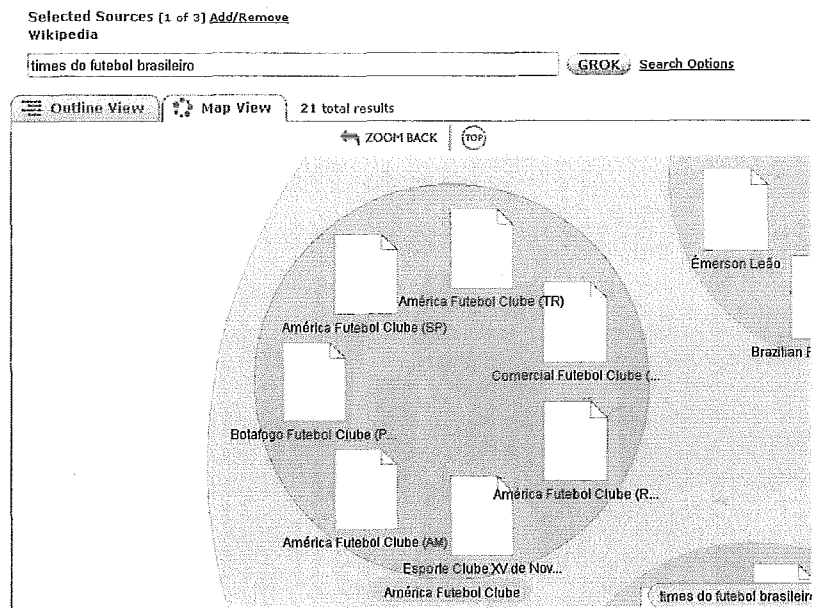


Figura 7 – Visualização exibida pelo sistema Grokker após selecionar o círculo América Futebol Clube (GROKKER, 2007).

O sistema GROKKER facilita a navegação do usuário através de hierarquias, porém se ele estiver representando hierarquias muito profundas o usuário terá que fazer muitas

operações de “zoom back” para retornar a visão geral do espaço de informação. A seguir, uma outra forma para representação de informação em espaço complexo é apresentada.

3.2.1.2. Mapas de Árvores (“TreeMaps”)

O *Mapa de Árvores* é explicada com maiores detalhes, pois foi utilizado na construção da ferramenta que implementa a abordagem proposta nesta dissertação. O *Mapa de Árvores* é uma técnica de visualização que permite exibir grandes conjuntos de informações hierárquicas através de métodos de preenchimento de espaço (SHNEIDERMAN, 1992). Isto ocorre pela divisão da área de exibição em uma seqüência aninhada de retângulos correspondentes a atributos de um conjunto de dados. O tamanho e a cor do retângulo podem ser utilizados para complementar a informação apresentada. As propriedades do *Mapa de Árvores* são enumeradas a seguir:

1. Existe somente sobreposição entre retângulos antecessores e predecessores.
2. A área do retângulo é proporcional ao seu tamanho.
3. O tamanho do retângulo é maior ou igual à soma do tamanho dos seus filhos.

A visualização provida pelo *Mapa de Árvores* permite uma mineração visual da informação, tornando mais fácil à descoberta de conhecimento a respeito dessa informação. As principais vantagens dessa técnica são mostradas a seguir:

1. Utilização eficiente da tela de visualização.
2. Preservação do contexto geral da informação.
3. Navegação rápida entre nós.
4. A estrutura do espaço de informação fica implícita no mapa exibido.
5. Muito útil na exibição de variáveis quantitativas dos dados.

Outro ingrediente chave na construção de *Mapa de Árvores* é a definição do algoritmo a ser utilizado na criação dos retângulos (como o estudo desses algoritmos não faz parte do escopo desta dissertação, somente uma visão geral sobre os mesmos será apresentada). Os mais conhecidos são cortar e fatiar, quadriculado, aglomerado, ordenado (pivô pelo meio e pivô por tamanho).

O *algoritmo de cortar e fatiar*, descrito em SHNEIDERMAN (1992), usa linhas paralelas para dividir o retângulo em retângulos menores que representam seus filhos. A cada nível da hierarquia a orientação das linhas (vertical e horizontal) é trocada. Apesar de

simples de se implementar, esse algoritmo frequentemente gera “layouts” que contêm muitos retângulos com alta taxa de aspecto (máximo valor da altura/comprimento e comprimento/altura), ou seja, muito longos e finos que podem ser difíceis de ver, selecionar, comparar em tamanho e rotular.

Quando as relações de aspecto das regiões retangulares são próximas de 1(um), os retângulos se aproximam de um quadrado e conseqüentemente o *Mapa de Árvores* consegue mostrar hierarquias maiores. Se houver nós com relações de aspecto muito pequenas, os retângulos podem desaparecer da tela, conforme mostrado na Figura 8.

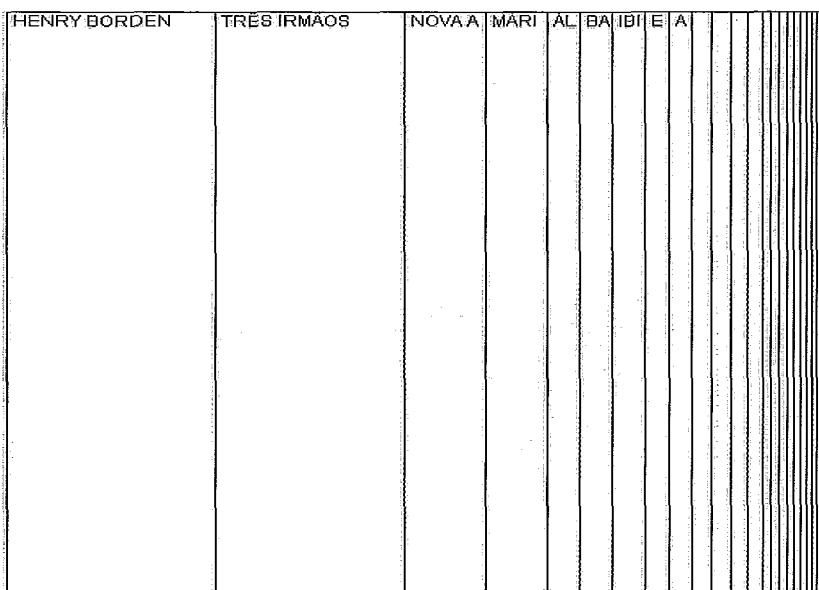


Figura 8 – Mapa de Árvore Cortar e Fatiar
(SHNEIDERMAN, 1992).

Alguns algoritmos criam visualizações mais claras da hierarquia com base no controle das relações de aspecto dos retângulos. Esses algoritmos melhoram a visualização de itens menores. Os Mapas de Árvores Aglomerados (“Cluster TreeMap”) (WATTENBERG, 1999) e Quadrículados (“Squarified TreeMap”) (BRULS *et al.*, 2000) usam algoritmos de recursão simples que reduzem substancialmente a relação de aspecto. Todavia, eles introduzem instabilidade na visualização a partir da mudança dinâmica dos dados e falham na preservação da ordenação dos dados. Isto é mostrado pelas diferenças

nos tons de cinza das Figura 9 e Figura 10. O algoritmo cortar e fatiar preserva a ordem dos dados enquanto que o aglomerado e quadriculado não.

Entretanto, uma alteração no conjunto de dados pode causar mudanças substanciais nos mapas produzidos por esses 2 algoritmos. Essas mudanças são indesejáveis quando os dados do *Mapa de Árvores* são atualizados muito freqüentemente, pois isto dificultaria a interpretação ou seleção de itens neste tipo de mapa.

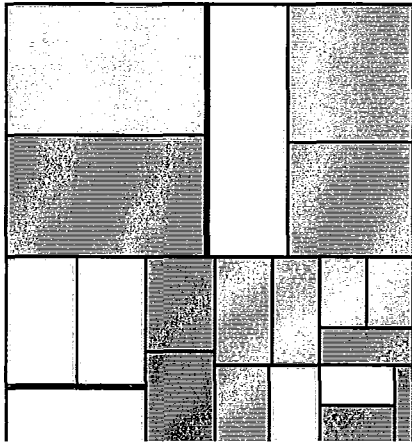


Figura 9 – Aglomerado (WATTENBERG, 1999).

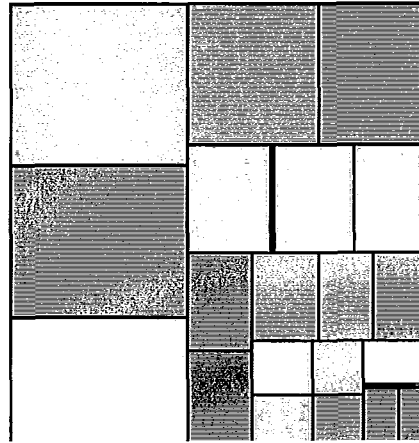


Figura 10 – Quadriculado (BRULS et al., 2000).

Objetivando contornar o problema da ordenação alfabética dos Mapas de Árvores em Aglomerado e Quadriculado, BENJAMIN et al. (2002) propõem a construção de Mapa de Árvores Ordenados através de dois algoritmos, pivô por tamanho (“pivot-by-size”) e pivô pelo meio (“pivot-by-middle”), estes produzem mapas com relações de aspecto que são menores que o cortar e fatiar, mas não tão bons como o aglomerado e quadriculado. Entretanto, os mapas gerados conseguem preservar razoavelmente a ordem em um cenário onde os elementos mudam freqüentemente. Logo, esses algoritmos podem ser uma boa escolha em situações onde a legibilidade, usabilidade e atualização freqüentes são itens importantes.

Ambos seguem um processo recursivo parecido, no qual começam com um retângulo R que será subdividido. A entrada é uma lista de itens ordenados por um índice e possuem uma área variada. Um item é escolhido como o pivô e colocado em um lado de R como um quadrado (relação de aspecto = 1) e os itens restantes colocados em três grandes retângulos que compõem o restante da área da tela. O algoritmo é aplicado recursivamente a cada um desses retângulos. No *pivô por tamanho* o item escolhido é o que possui a maior

área. No *pivô pelo meio* o item do meio da lista é escolhido, ou seja, se uma lista possui N itens o item escolhido é o item com número $n/2$. A Figura 11 ilustra um exemplo de Mapa de Árvore gerado por esses dois algoritmos respectivamente.

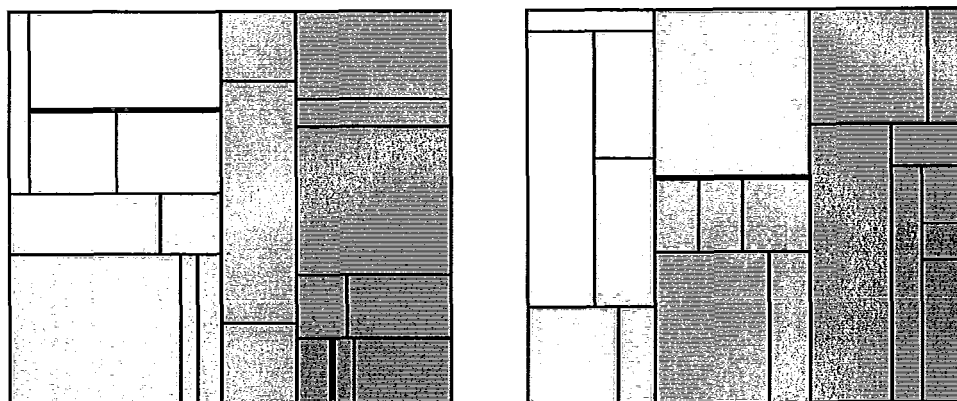


Figura 11 – Pivô por Tamanho e Pivô pelo Meio (BENJAMIN et al., 2002).

Um exemplo de utilização da técnica Mapa de Árvores pode ser vista em ADAM & MARC (2006), conforme mostrado na Figura 12. Neste trabalho, essa técnica é utilizada na visualização dos “e-mails” de um usuário, com objetivo de exibir a hierarquia de contatos a partir do domínio implícito em seus “e-mails”. No Mapa de Árvores da Figura 12, o tamanho do retângulo é proporcional ao número de mensagens de um contato específico. Já a cor indica o número de mensagens não lidas.

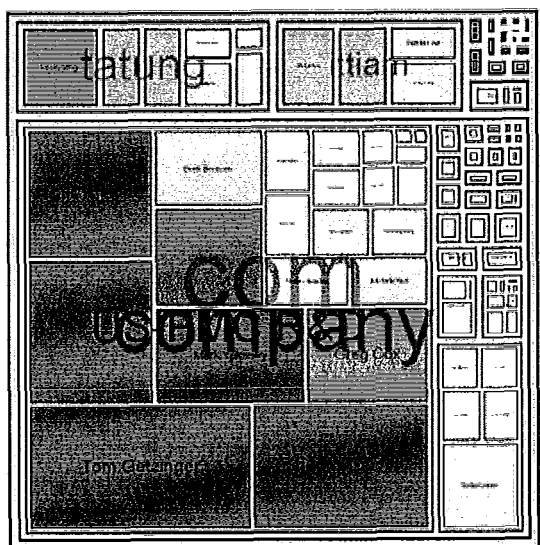


Figura 12 – Mapa de Árvore que representa a caixa de mensagens de um usuário (ADAM & MARC, 2006).

3.2.1.3.Árvore Hiperbólica (Hyperbolic Viewers)

A geometria hiperbólica é outra técnica muito utilizada na visualização de informações hierarquizadas. Essa técnica produz uma visão “fish-eye” (olho de peixe) da árvore hierárquica, permitindo que o usuário visualize a árvore inteira e os respectivos detalhes de uma certa área em uma única tela (LAMPING & RAO, 1995). Para se obter uma visão mais detalhada de uma determinada área, basta o usuário arrastar um nó desta área, em contra partida, os nós que não pertencem à área desejada ficaram encolhidos no canto. A interatividade provida por essa técnica facilita a busca em grandes hierarquias. O ponto negativo dessa técnica, assim como das outras estratégias utilizadas para se visualizar espaços complexos, é que apesar do caminho topológico da informação ser mostrado, o relacionamento semântico dos nós da árvore é negligenciado.

Um exemplo do uso de árvore hiperbólica pode ser visto no GCC (OLIVEIRA et al., 2005), no qual esta árvore é utilizada na visualização da hierarquia dos conceitos definidos pela Árvore de Conhecimento do CNPq (CNPq, 2007) , conforme mostrado na Figura 13.

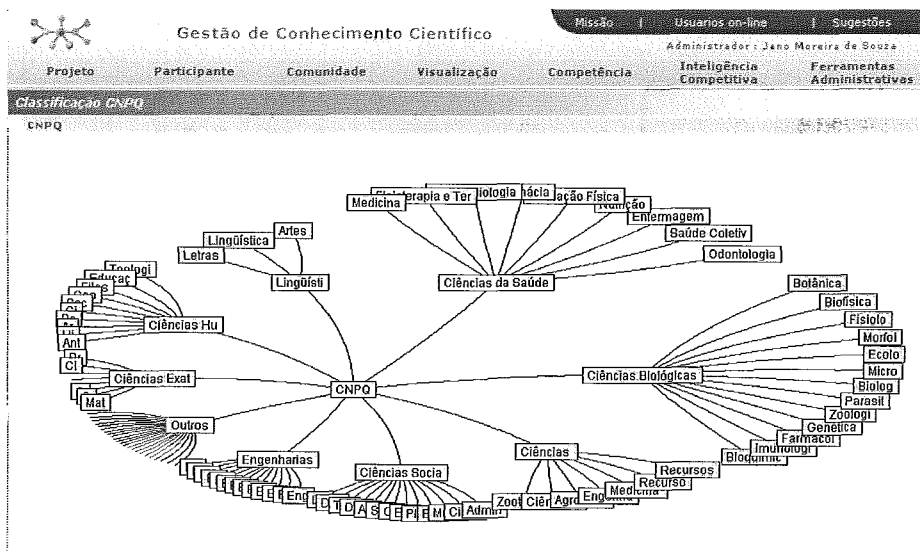


Figura 13 – Árvore CNPq, visão hiperbólica provida pelo GCC (OLIVEIRA et al., 2005).

A árvore Bifocal é outro exemplo do uso da visualização hiperbólica, vide Figura 14. Essa técnica também se baseia no uso do conceito foco e contexto, porém utiliza dois

focos e não apenas um. A árvore é dividida em dois sub-diagramas conectados, onde o da direita exibe a informação detalhada e o da esquerda mantém uma perspectiva geral do espaço de informação. Isto possibilita ao usuário obter uma visão detalhada da área desejada, sem que ele perca a percepção geral do contexto.

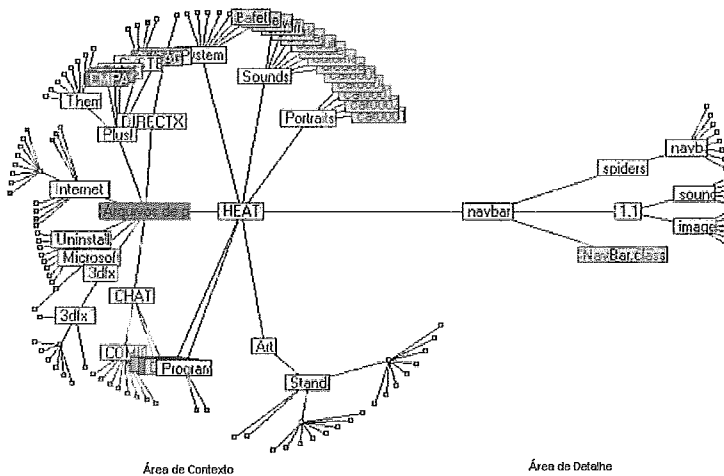


Figura 14 – Árvore Bifocal. Adaptada de CAVA et al. (2002).

3.2.2. Sistemas de Visualização de Rotas

Quando o principal objetivo do usuário é visualizar a topologia da informação e não navegar através da mesma, os *sistemas de visualização de rotas* podem ser utilizados. Esses sistemas são muito utilizados na visualização de grandes espaços de informação, como por exemplo: a distribuição de servidores da Internet ou caminho entre hipertextos. Geralmente esses sistemas não dão uma visão detalhada de elementos individuais, mas tentam mostrar padrões na visão geral. Devido ao enorme volume de dados, a visualização provida pode se tornar densa e extrapolar o limite cognitivo do usuário. A Figura 15 mostra o Opte Project (OPTE PROJECT, 2007), que é um sistema de rota utilizado para visualização da distribuição de servidores da Internet.

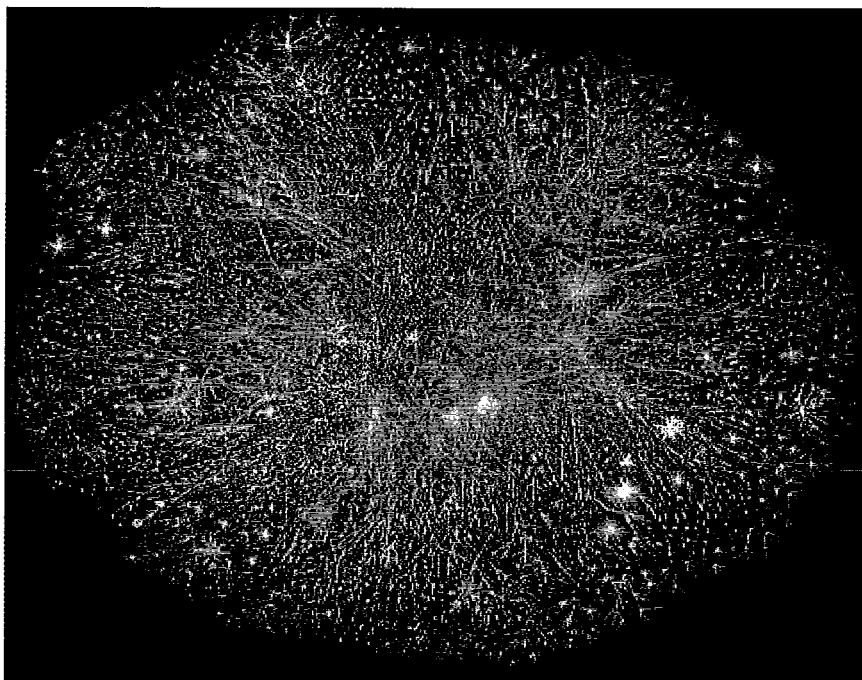


Figura 15 – Mapa produzido pela Opte Project em novembro de 2003 (OPTE PROJECT, 2007).

Os sistemas para visualização de *espaço de informação complexo* são muito úteis na visualização da hierárquica ou topologia da informação. Entretanto, estes sistemas mostram-se ineficientes quando o objetivo do usuário é analisar o relacionamento semântico da informação. Para solucionar este problema, são utilizados os *sistemas de visualização de contexto* apresentados a seguir.

3.3. Sistemas de Visualização de Contextos

Interfaces de informação modernas, especialmente as de busca por palavra-chave, retornam muitos resultados, porém provêem pouco “insight” no relacionamento semântico dos objetos digitais retornados. Este fato é decorrente dos resultados serem apresentados em uma dimensão ao usuário, isto é, sem relacionamento e classificação. Por exemplo, os resultados retornados por uma busca executada no GOOGLE (GOOGLE, 2007). Portanto, o grande desafio no projeto de visualizações de contexto é ilustrar, através da criação de uma infra-estrutura teórica e tecnológica que dê suporte à “web” semântica, os objetos que possuem uma relação multidimensional com outros (BERNERS *et al.*, 2001).

Nesta seção, os exemplos de sistemas de visualização de contexto estão divididos em duas sub-categorias:

1. *Sistemas de exploração* (“Explore Systems”): são sistemas de visualização interativos, que permitem a manipulação exploratória da visualização.
2. *Sistemas de mapa* (“Chart Systems”): são sistemas de visualização que representam o espaço de informação através de mapas estáticos, porém navegáveis.

3.3.1. Sistemas de Exploração

Sistemas de exploração são utilizados na visualização do conteúdo semântico da informação, permitindo que o usuário explore o espaço semântico através de mudanças interativas na distribuição espacial dos “clusters”. Essas mudanças geralmente são decorrentes da inserção de palavras no mapa da visualização, com isso, os elementos do mapa podem se aproximar ou afastar dessas palavras conforme suas similaridades com as mesmas. Logo, a inserção de palavras permite ao usuário alterar a aparência do mapa de exibição de forma intuitiva e interativa.

Esta seção está dividida em duas categorias de visualizações: as que se baseiam na *relevância semântica* (“semantic relevance”) e na *consulta dinâmica* (“dynamic query”).

3.3.1.1. Relevância Semântica

O grande problema de se organizar objetos de um arquivo em uma ontologia⁴ hierárquica, é que estes objetos freqüentemente pertencem a múltiplas classes. Para minimizar esse problema muitas ferramentas utilizam a técnica denominada “Cluster Map”, a qual procura classificar a informação com base em seus conceitos ontológicos (FLUIT *et al.*, 2005).

O AutoFocos Personal (AUTOFOCOS PERSONAL, 2007) é um exemplo de aplicativo que faz uso dessa técnica, conforme mostrado na Figura 16. Neste programa, os documentos são representados por círculos coloridos e agrupados num grande círculo

⁴ Ontologia é uma teoria sobre um domínio que especifica um vocabulário de entidades, classes, propriedades, predicados e funções; e um conjunto de relações que necessariamente interliga esse vocabulário. (OLIVEIRA, 1999).

sombreado (“cluster”), com outros documentos que se relacionam à mesma palavra-chave. Se uma segunda palavra-chave for inserida, uma nova visualização é gerada, na qual os documentos relacionados à primeira palavra-chave ficam acima, os relacionados a segunda ficam abaixo e os relacionados a ambas palavras ficam centralizados. O processo se repete para cada nova palavra-chave inserida.

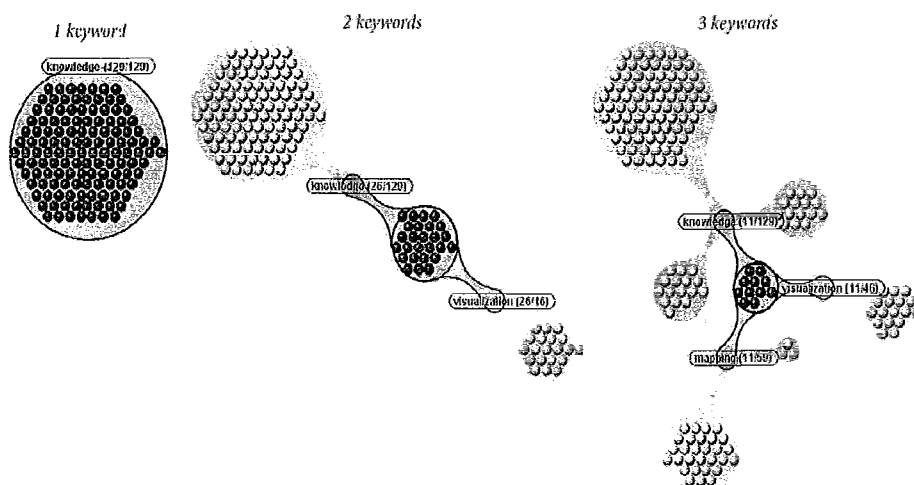


Figura 16 – Programa AutoFocus Personal (AUTOFOCOS PERSONAL, 2007).

O sistema Sinnzeug (SINNZEUG, 2007) é outro exemplo de utilização da técnica “Cluster Map”. Este programa utiliza o modelo de “datacloud” magnético numa interface de busca que lida com 25 “sites”. Cada “site” é representado por um pequeno círculo preto distribuído sobre um mapa, conforme mostrado na Figura 17. Os “sites” são atraídos pelas palavras-chave inseridas na tela, de forma que os “sites” que possuem conteúdo semântico similar fiquem agrupados. Posicionando-se o mouse sobre algum dos pontos, uma descrição do “site” é exibida.



Figura 17 – Programa Sinnzeug (SINNZEUG, 2007).

Estas visualizações não foram utilizadas na abordagem proposta nesta dissertação. Entretanto, a idéia de aproximar elementos, conforme sua similaridade, foi aplicada na visualização que compara a similaridade de perfil MBTI dos pesquisadores, que será apresentada na Seção 4.3.3.1.

A próxima seção descreve a técnica *consulta dinâmica* (SHNEIDERMAN, 1994) que também é muito utilizada nas visualizações relacionadas à exploração do aspecto semântico da informação.

3.3.1.2. Consultas Dinâmicas

A técnica de consultas dinâmicas (SHNEIDERMAN, 1994) usa manipulação direta para possibilitar que usuários controlem parâmetros visuais de consulta, gerando rapidamente uma representação visual dos resultados de uma busca na base de dados consultada. Dentre as diversas vantagens dessa técnica, destacam-se:

1. Facilidade de aprendizagem do usuário sobre como efetuar uma consulta, sem precisar entender complexas sintaxes de linhas de comando (característica derivada do modo de manipulação direta).

2. Minimização das possibilidades de erros do usuário com relação à sintaxe da consulta expressa pelos estados dos controles visuais de parâmetros.
3. Rapidez entre a formulação da consulta e a exibição do seu resultado, o que pode facilitar ao usuário construir um modelo interno ou mapa cognitivo dos dados sob análise (SPENCE, 2001).

O programa “Film Finder” (AHLBERG & SHNEIDERMAN, 1994), é um exemplo de aplicação que utiliza consultas dinâmicas para permitir a exploração de uma base de dados de filmes. Como mostrado na Figura 18, esse programa apresenta um gráfico de dispersão, no qual cada retângulo é uma marca que representa um filme. A cor de cada marca indica o tipo de filme a que ela se refere, segundo uma legenda apresentada na parte inferior do gráfico. A posição das marcas na vertical indica os anos de lançamento dos filmes, enquanto a posição das marcas na horizontal indica a popularidade desses filmes. As consultas dinâmicas são executadas através dos controles que permitem escolher título, ator, atriz, diretor, tamanho e classificação dos filmes a serem exibidos no gráfico.

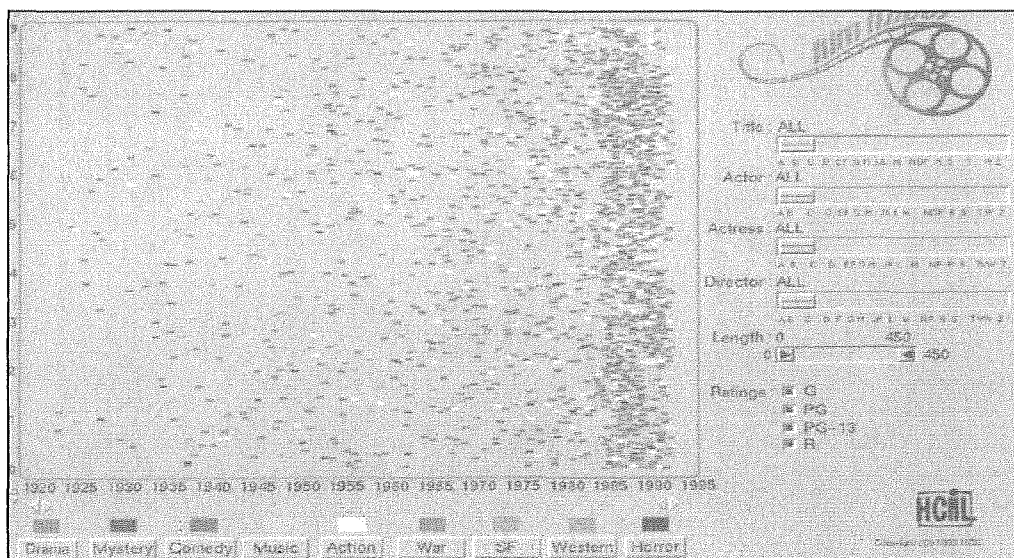


Figura 18 – Programa Film Finder (AHLBERG & SHNEIDERMAN, 1994).

Os sistemas de Mapas é outro conjunto de visualizações muito utilizado na exploração do conteúdo semântico da informação. A diferença é que os sistemas deste conjunto geram visualizações estáticas, porém navegáveis. A próxima seção descreve estes sistemas com maiores detalhes.

3.3.2. Sistemas de Mapa

Estas visualizações ilustram o relacionamento contextual entre objetos relacionais. *Mapa* refere-se a uma metáfora geográfica usada por muitos exemplos, produzindo “landscape” (2D ou 3D) do espaço de informação que está sendo representado (HAREL, 1987). Estes mapas são interativos e geralmente permitem a reposição e “zooming” no “landscape” para se visualizar mais detalhes de um tópico em particular. Nos sistemas de Mapa, a distribuição espacial dos “clusters” semânticos são estáticas e pré-geradas das estatísticas da análise semântica, diferentemente dos sistemas de exploração que são geradas em tempo real.

Esta seção está dividida em três categorias: *máquinas de busca na Internet* (“internet search engines”), *metáforas de dados* (“data landscapes”) e *ambientes virtuais* (“virtual environments”).

3.3.2.1. Máquinas de Busca na Internet

Nas máquinas de buscas atuais (GOOGLE, YAHOO etc), os resultados são apresentados como uma *lista ranquedada* com base em um algoritmo de indexação. As buscas realizadas por essas máquinas retornam muitos resultados irrelevantes, pois são realizadas com base em palavras-chave. Uma redução na quantidade de resultados irrelevantes retornado pode ser conseguida através do aumento da indexação semântica dos resultados buscados.

Vivisimo (Vivisimo, 2007) é um programa de mineração de dados e análise de texto. Seu algoritmo gera “clusters” de documentos texto, classificando-os em categorias. Apesar da interface do Vivisimo ser baseada em texto, a classificação semântica dos resultados facilita a busca de informação, conforme ilustrado na Figura 19.



Figura 19 – Exemplo do resultado retornado por uma busca realizada no programa Vivissimo (VIVISSIMO, 2007).

O KARTOO (KARTOO, 2007) utiliza um algoritmo de clusterização similar ao do VIVISSIMO. No entanto, os resultados retornados pelo KARTOO são exibidos em categorias semânticas majoritárias e apresentados na forma de uma visualização cartográfica interativa, conforme mostrado na Figura 20. Este aplicativo também utiliza um sistema de ícones para diferenciar os tipos e tamanho dos documentos. Um simples clique sobre as palavras presentes no mapa permite que estas passem a fazer parte da busca e conseqüente uma nova visualização é gerada. Os pontos negativos desta ferramenta são o seu elevado grau de aprendizado e o número de resultados exibidos por vez, variando entre 10 e 15.

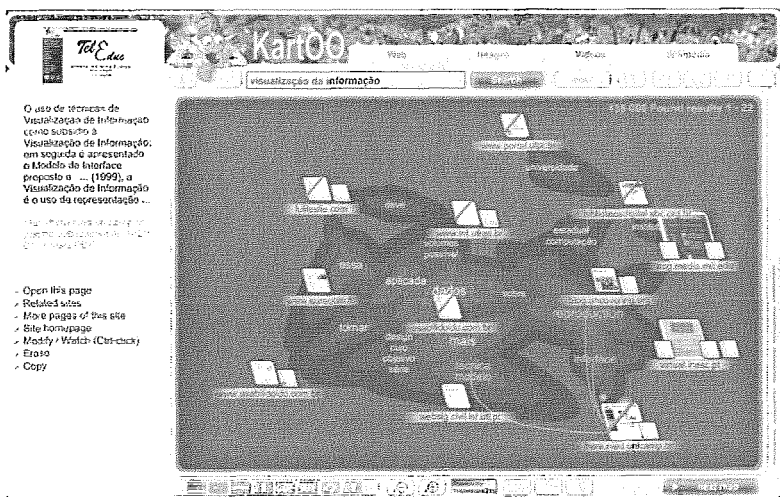


Figura 20 – Exemplo do resultado retornado pela busca realizada no KARTOO (KARTOO, 2007).

Esta seção apresentou dois exemplos de máquinas de busca que procuram adicionar algum aspecto semântico aos seus resultados. Na próxima seção a segunda categoria de sistema de Mapas é apresentada.

3.3.2.2. Metáfora de Dados

As visualizações pertencentes a este grupo são chamadas *sistemas de informação geográfica*. Estas visualizações exibem os dados em espaços geográficos, utilizando a facilidade das pessoas de interpretar mapas para que estas entendam a informação com um menor esforço cognitivo.

O algoritmo de *mapas auto-organizados* (KOHONEN, 1989) foi desenvolvido para organizar milhares de arquivos com base em suas similaridades semânticas. Esse algoritmo produz mapas 2D onde documentos relacionados aparecem próximos. Muitos sistemas de informação geográfica, dentre os quais são apresentados o WebSOM (WEBSOM, 2007) e o VXINSIGHT (KRISHNAN *et al.*, 2007), fazem uso desse algoritmo para organizar a informação a ser exibida ao usuário.

WebSOM é um navegador visual utilizado para organizar discussão de grupos de usuários na Internet. Neste navegador, os tópicos das discussões são rotulados e coloridos (a cor é gerada com base na quantidade de informações de um tema específico). Um “zoom” neste mapa pode ser feito clicando-se sobre ele, após dois clique, pontos brancos são apresentados. Esses pontos representam os assuntos debatidos, ao selecioná-los, as mensagens referentes ao assunto são exibidas em uma janela. Em algumas seções deste mapa, a lógica do agrupamento semântico não é evidente, pois temas não relacionados estão próximos entre si (por exemplo, música é rodeada por “linux”). O exemplo da Figura 21 foi construído a partir de imagens copiadas de uma demonstração “on-line” desta ferramenta.

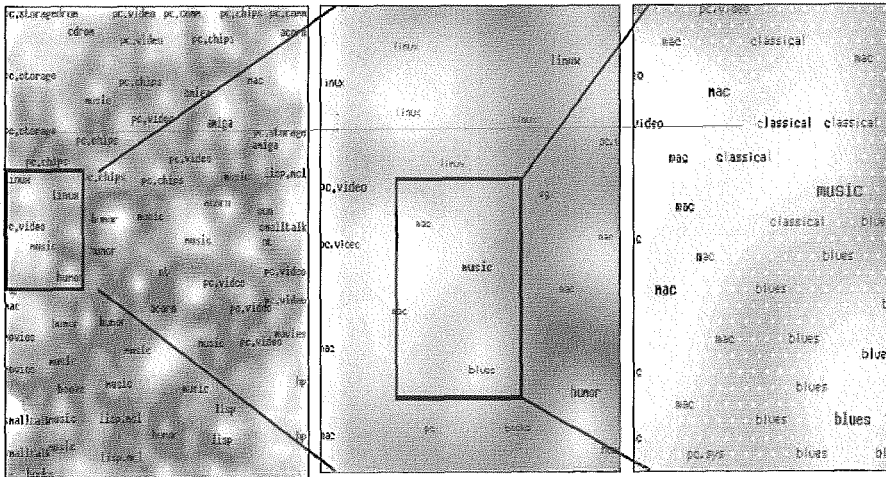


Figura 21 – Exemplo da ferramenta WebSOM (WEBSOM, 2007).

O VXINSIGHT é um aplicativo utilizado para visualizar os dados de um grande banco de dados, conforme apresentado na Figura 22. Este programa exibe os resultados da busca como “landscape” de montanhas 3D. Cada montanha representa um conjunto de documentos de uma área particular, sendo que tamanho de cada montanha varia conforme o número de documentos em cada conjunto. Outra característica interessante deste sistema é a definição de um parâmetro de tempo que permite a variação do tamanho dos “landscapes” conforme o período de tempo desejado. Outras utilizações desse aplicativo podem ser encontradas em “PNL” (PNL, 2007).

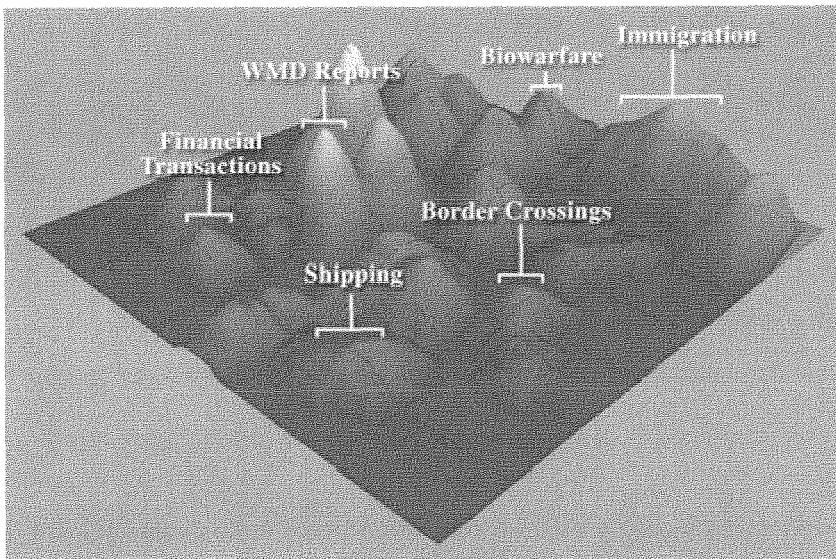


Figura 22 – Exemplo de utilização do VXINSIGHT (KRISHNAN et al., 2007).

A próxima seção apresenta a última categoria de sistema de Mapas que são os Ambientes virtuais.

3.3.2.3. Ambientes Virtuais

Uma área de pesquisa promissora é a utilização de visualizações navegáveis em 3D na exibição de dados. Ao se transformar os mapas de informação de planos 2D para 3D, aumenta-se o grau de interatividade e complexidade da visualização. O aumento da interatividade é decorrente do fato um ambiente 3D ser mais propício à colaboração e a ações simultâneas de seus usuários, por exemplo, jogos de computadores 3D na Internet. Entretanto fornecer e manter essas ações e interatividades coerentes envolve grande complexidade.

CHEN (1999) discute algumas das principais questões da integração da realidade virtual com informações e apresenta algumas aplicações-chave. Pelo fato desta dissertação não utilizar visualizações 3D, este assunto não será aprofundado e nem exemplificado.

Muitas vezes, os usuários estão interessados no fluxo da informação e não em seu aspecto semântico ou arquitetural o que torna as visualizações estudadas até agora ineficazes. Logo, um outro grupo de sistemas direcionado a atender essa necessidade se faz necessário, o qual é apresentado a seguir.

3.4. Visualização de Espaços Dinâmicos

Nas estruturas culturais modernas, houve um grande aumento de fluxo nas esferas tecnológicas, políticas, ideológicas e econômicas (CAMPANHOLA & SILVA, 2000). Logo, os sistemas de visualização de espaço dinâmico procuram ilustrar esses movimentos, fluxos e mudanças no tempo ou espaço de modo que o usuário consiga avaliá-los adequadamente.

Nesta seção, os exemplos de sistemas dinâmicos estão divididos em duas subcategorias:

1. *Fluxo* (“flow”): sistemas de visualização que ilustram a quantidade, taxa e caminho do movimento de alguma entidade.

2. *Evolução* (“evolve”): sistemas de visualização que ilustram o caminho de influência ou mudança nos sistemas com o passar do tempo.

3.4.1. Sistemas de Visualização de Fluxo

Sistemas de visualização de fluxo são utilizados para representar a quantidade, taxa e caminho do movimento de alguma entidade. Para atingir este objetivo, esses sistemas utilizam vetores com direções, espessuras e cores diferentes. Uma característica peculiar dos sistemas de visualização de fluxo é a pouca ou nenhuma interatividade das visualizações produzidas. A seguir são apresentados dois exemplos de sistema de fluxo. O primeiro exemplo, que pode ser visto na Figura 23, utiliza *diagramas de arcos* (WATTENBERG, 2002) para representar trechos similares em músicas, onde a espessura da linha indica o tamanho do trecho que se repete. Enquanto o segundo exemplo, conforme mostrado na Figura 24, ilustra uma visualização gerada pelo programa “SeeNet3D” (EICK *et al.*, 1996) para representar o fluxo de comunicação entre diferentes espaços geográficos. Repare que neste exemplo, diferentemente do primeiro, a quantidade de fluxo é ilustrada pela cor das setas e não por sua espessura.

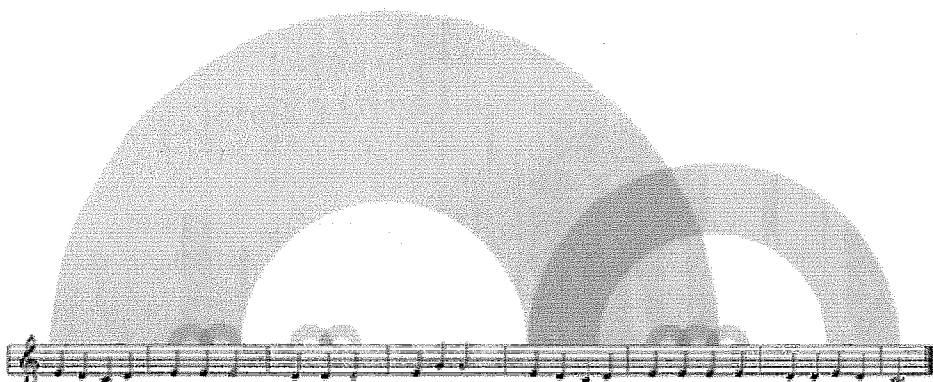


Figura 23 – Exemplo do arco de diagramas (WATTENBERG, 2002).

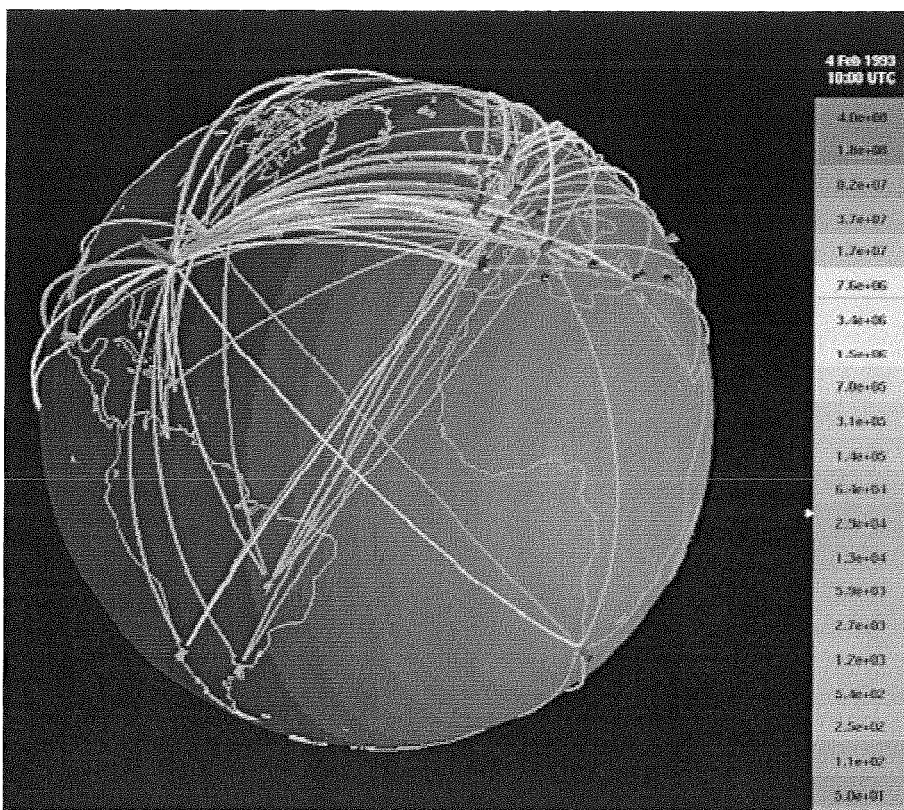


Figura 24 – Exemplo de visualização gerada pelo programa SeeNet3D (EICK et al., 1996).

Estas visualizações não são utilizadas na abordagem proposta nesta dissertação. A idéia de representar variações de taxas das entidades foi aplicada para representar o quanto um especialista aprendeu um determinado assunto com o passar do tempo, conforme explicado na seção 4.3.3.2.

Um outro conjunto de *visualização de fluxo* que objetiva avaliar as alterações nos sistemas a partir dos eventos-chave ou mudanças ao passar do tempo é apresentado a seguir.

3.4.2. Sistemas de Evolução

Sistemas de Evolução são utilizados para representar os principais eventos ou mudanças que ocorreram num sistema com o passar do tempo. Para alcançar este objetivo,

esses sistemas fazem uso de eixos temporais em suas representações. A seguir são apresentados dois exemplos de Sistema de Evolução.

O primeiro exemplo, conforme mostrado na Figura 25, refere-se ao aplicativo “ThemeRiver” (HAVRE et al., 1999), que foi desenvolvido para ajudar seus usuários a reconhecerem tendências em um conjunto de documentos durante um período de tempo. Cada linha representa um tema, a espessura de cada linha aumenta conforme o número de documentos relacionados ao seu tema aumenta.

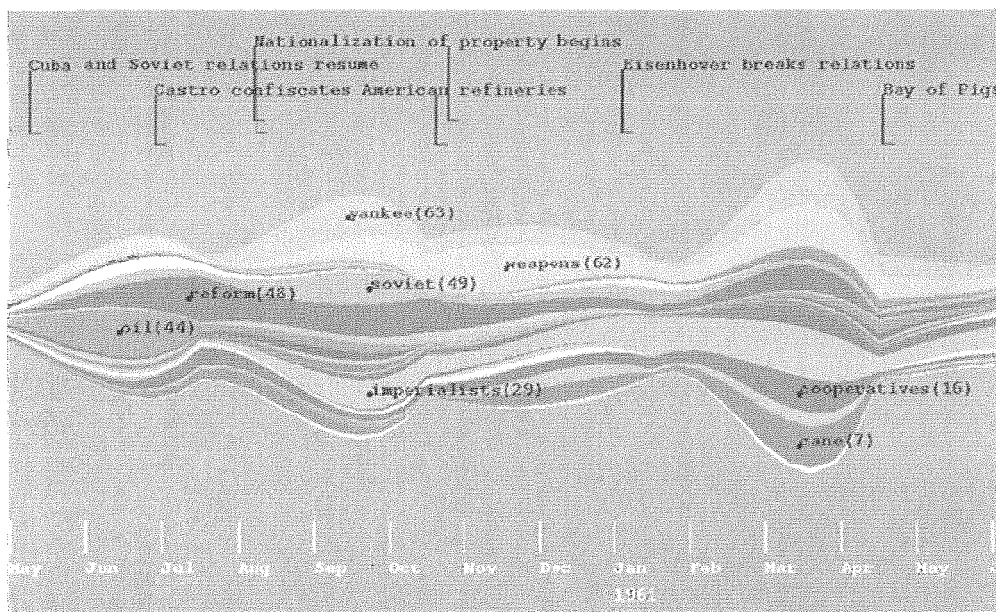


Figura 25 – Imagem de uma visualização gerada pelo aplicativo ThemeRiver (HAVRE et al., 1999).

O segundo exemplo, ilustrado na Figura 26, refere-se ao aplicativo REVISIONIST (REVISIONIST, 2007), produzido com intuito de facilitar a visualização de como o código de um sistema evoluiu. Neste, cada versão do código é listada numa coluna, as quais são conectadas por curvas coloridas nos locais onde ocorreram mudanças. Quanto maior for a número de mudanças, mais espessa fica a curva de conexão.

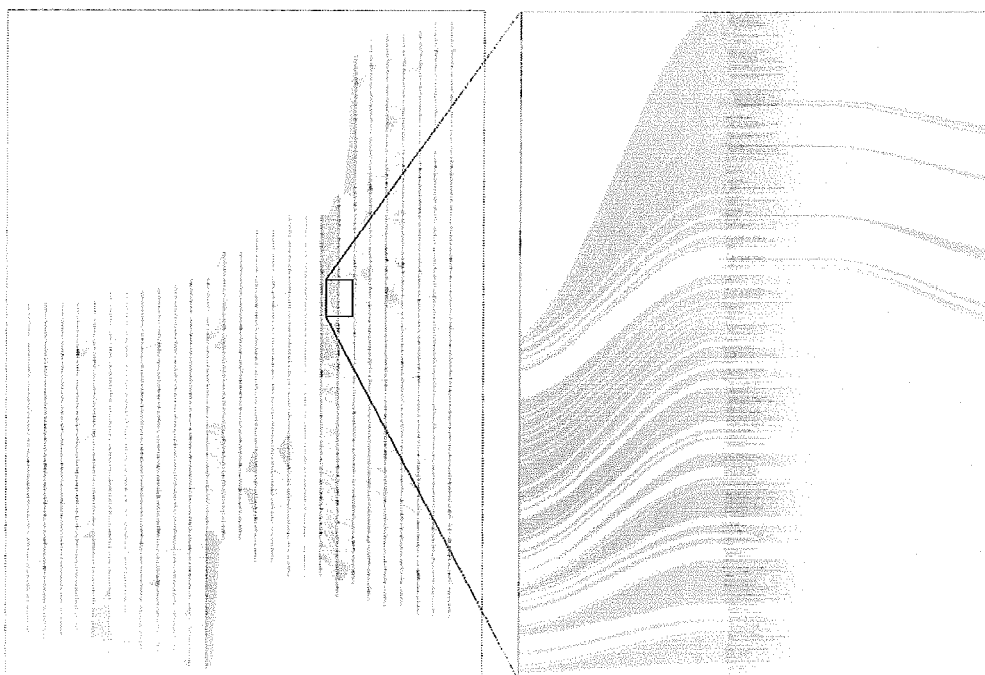


Figura 26 – Imagem de uma visualização gerada pelo Revisionist (REVISIONIST, 2007).

Este foi o último exemplo referente ao estado da arte de Visualização da Informação. Este estudo apresentou uma visão geral sobre sistemas e técnicas de Visualização de Informação com base na classificação proposta por JUDELMAN (2004).

Uma outra proposta de classificação pode ser vista em LENGLER & EPPLER (2007), no qual métodos de visualização são apresentados através de uma tabela periódica, conforme mostrado na Figura 27. Um método de visualização é uma representação gráfica, sistemática, externa, baseada em regras e permanente que apresenta as informações de uma forma que propicia a aquisição de conhecimentos, o desenvolvimento de uma concepção elaborada e a transmissão de experiências (LENGLER & EPPLER, 2007). De acordo com esses autores, a tabela periódica de visualização não deve ser considerada somente um repositório de visualizações, mas sim uma abordagem inicial de como organizar, onde aplicar e como integrar o grande número de visualizações existentes. Nessa tabela os métodos de visualização são classificados em cinco dimensões:

1. Complexidade da visualização: alta ou baixa, refere-se ao número de interdependências entre os elementos a serem visualizados.

2. Área de aplicação: visualização de dados, informação, conceito, metáfora, estratégia e conhecimento composto.
3. Ponto de visão: detalhe (destaca itens individuais), visão geral (visualização do todo), detalhe + visão geral (ambas ao mesmo tempo).
4. Tipo de ajuda esperada: convergente (redução da complexidade através de análise e síntese) e divergente (adição de complexidade objetivando a obtenção de “insight”).
5. Tipo de informação representada: estruturada (por exemplo: redes e hierarquias) e processos (com fases em tempos cíclicos ou contínuos seqüenciais).

Assim como na tabela periódica de química, a tabela de métodos de visualização é construída em duas dimensões: períodos (correspondem a uma linha da tabela) e grupos (correspondem a uma coluna da tabela). Nesta tabela, o *período* representa a complexidade da visualização: quanto mais à direita uma visualização, maior é sua complexidade. A cor das letras de cada elemento da tabela é utilizada para ilustrar a dimensão *tipo de informação representada*, sendo a cor *preta* para a informação estruturada e *azul* para a de processo. O ponto de visão (detalhe [■], visão geral [☼] e visão geral com detalhe [⊕]) e o tipo de ajuda esperada (convergente[<>] e divergente [><]) são apresentados acima da sigla do elemento de visualização. Por fim, a dimensão *área de aplicação* é representada através dos grupos, por exemplo: no grupo quatorze (coluna 14) estão somente visualizações aplicadas na área de gerenciamento de projetos. Além de também estar subdividida em seis categorias que são diferenciadas pela cor de fundo:

1. Visualização de Dados (amarelo) – representações visuais esquemáticas dos dados quantitativos (através de eixos ou não).
2. Visualização de Informação (verde) – uso de representações visuais interativas para ampliar a cognição.
3. Visualização de Conceito (verde limo) – métodos para elaborar conceitos, idéias, planos e análises.
4. Visualização de Metáfora (azul claro) – Metáforas visuais posicionam a informação graficamente com objetivo de organizá-la e estruturá-la. Elas também possibilitam a geração de “insight” através das principais características da metáfora que está sendo utilizada.

5. Visualização de Estratégia (violeta) – refere-se à utilização representações visuais complementares para melhorar a análise, desenvolvimento, formulação, comunicação e execução de estratégias nas organizações.
6. Visualização Composto (roxo) – o uso complementar de diferentes formas de representação gráfica de uma única vez.

C cosidera	Data Visualization Visual representations of quantitative data in schematic form (either with or without axes)										Strategy Visualization The systematic use of complementary visual representations in the analysis, development, formulation, communication, and implementation of strategies in organizations.										G graphical information	
Tb table	Ca cartesian coordinates	Information Visualization The use of interactive visual representations of data to amplify cognition. This means that the data is transformed into an image, it is mapped to screen space. The image can be changed by users as they proceed working with it.										Metaphor Visualization Visual Metaphors position information graphically to organize and structure information. They also convey an insight about the represented information through the key characteristics of the metaphor that is employed.										H hierarchy
Pi pie chart	L line chart	Concept Visualization Methods to elaborate (mostly) qualitative concepts, ideas, plans, and analyses.										Compound Visualization The complementary use of different graphic representation formats in one single schema or frame.										Cc critique
B bar chart	Ae area chart	R radar chart	Pa panda coordinates	Hy hyperbolic tree	Cy cycle diagram	T timekete	Ve veva diagram	Mi mindmap	Sq square of opposition	Ce concentric circles	Ar argument table	Sw swim lane diagram	Gc gantt chart	Pm projective diagram	D diamond diagram	Pr partition tree	Kn knowledge map					
Hi histogram	Se scatterplot	Sa sashy diagram	In information hour	E entity relationship diagram	Pt petri net	Fl flow chart	Cl clustering	Le layer chart	Py pyramid	Ce concentric circles	Ti twinning map	Dt decision tree	Cp span critical path method	Cf concept fan	Co concept map	Ic iceberg	Lm learning map					
Tk tubey box plot	Sp spectrogram	Da data map	Tp treemap	Cn code tree	Sy system dyn./simulation	Df data flow diagram	Se semantic network	So soft system modeling	Sn strategy map	Fo force field diagram	Ib the argumentation map	Pr process event chains	Pe petri chart	Ev evolution knowledge map	V vse diagram	Hh hearsen's led chart	I informational					

Cy Process Visualization
Hy Structure Visualization
☀ Overview
□ Detail
☉ Detail AND Overview
< > Divergent thinking
> < Convergent thinking

Note: Depending on your location and connection speed it can take some time to load a pop-up picture.
© Ralph Lengler & Martin J. Eppler, www.visual-literacy.org
version 1.5

Su supply demand axes	Pe performance charting	St strategy map	Oc organization chart	Ho house of quality	Fd feedback diagram	Ft future tree	Mq magic quadrant	Ld life-cycle diagram	Po porter's five forces	S s-cycle	Sm stakeholder map	Is ishikawa diagram	Tc technology roadmap
Ed edgeworth box	Pf portfolio diagram	Sg strategic game board	Mz maltzberg's orgnigraph	Z zandy's morphological box	Ad affinity diagram	De decision discovery diagram	Bm bcg matrix	Stc strategy canvas	Vc value chain	Hy hype cycle	Sr stakeholder rating map	Ta taps	Sd spray diagram

Figura 27 – Tabela periódica de métodos de visualização (LENGLER & EPPLER, 2007).

Apesar da proposta de LENGLER & EPPLER (2007) ser muito interessante, nós optamos por utilizar a classificação de JUDELMAN (2004) pelo fato desta categorizar as visualizações dentro de três espaços de informações (contexto, complexo e dinâmico), os quais estavam condizentes com as idéias da abordagem proposta, facilitando assim a escolha de quais visualizações utilizar.

Como o objetivo dessa dissertação é produzir uma abordagem que utilize técnicas de Visualização da Informação para recomendar pesquisadores substitutos no cenário

científico, um estudo sobre o estado da arte referente a sistemas de recomendação se fez necessário. Este estudo é apresentado a seguir.

3.5. Trabalhos Relacionados sobre Sistema de Recomendação

Hoje em dia, é cada vez maior a quantidade de informação a ser analisada para se encontrar uma determinada pessoa ou especialista. De acordo com SEID & KOBASA (2003), existem duas razões principais para se buscar especialistas: quando eles são fontes de informação ou são as pessoas aptas a executar determinadas tarefas ou funções social. Logo um sistema que processe toda essa informação e detecte uma pessoa, segundo algum critério, trará um grande benefício a esse processo de busca.

Segundo REICHLING *et al.*, (2005), um sistema de recomendação de especialistas (SRE) retorna referências para atores humanos que são identificados como “especialistas” no domínio requisitado. De acordo com TERVEEN & MCDONALD (2005), alguns autores estão denominando esses sistemas de recomendação como *Sistemas de Combinação Social* (SCS), “Social Matching Systems”, onde a principal diferença é que os sistemas de recomendação tradicionais, que indicam objetos, procuram resolver o problema da sobrecarga de informação, enquanto os SCS revelam informações sobre os usuários (CRUZ *et al.*, 2007) de maneira a auxiliar na socialização. Ou seja, um SRE pode recomendar coisas (como materiais, por exemplo), como pessoas para a formação de comunidades ou equipes. Um SRE é um SCS, mas não necessariamente um SCS é um SER. Segundo BONHARD *et al.* (2007), sistemas de recomendação de itens e pessoas não devem ser mutuamente excludente, pois podem se beneficiar um do outro, já que estão diretamente relacionados. Como a ferramenta proposta nessa dissertação procura recomendar substitutos a partir do processamento de seu conjunto de informações, podemos considerar essa ferramenta um SRE ou SCS. Iremos adotar a denominação SRE por considerarmos este nome mais intuitivo.

Durante o levantamento da literatura foi encontrado um grande número de SREs. Entretanto, nenhum desses sistemas utiliza diferentes técnicas de visualização da informação para representar cada faceta do perfil dos especialistas, e nem são direcionados

a substituição no meio acadêmico. Apesar disso, esses sistemas são considerados trabalhos correlatos, pois possuem o mesmo objetivo da proposta dessa dissertação que é recomendar uma pessoa a partir do seu perfil. A seguir são citados alguns dos sistemas encontrados, que foram escolhidos com base na diferença dos critérios ou técnicas utilizados na indicação dos especialistas.

EXPERT FINDER (MATTOX *et al.*, 1999), utilizado na empresa MITRE⁵ (MITRE, 2008), primeiramente busca documentos relevantes nas fontes de informações disponíveis (relatórios técnicos, resumos, notícias e páginas pessoais) que contém os termos utilizados na consulta de busca. Os documentos retornados são categorizados em dois grupos, documentos que um funcionário publicou a respeito de um tópico e documentos que mencionam funcionários relacionados a um tópico. No caso dos documentos de autoria própria, o sistema usa o número de documentos publicados pelo funcionário no cálculo do peso. No segundo grupo de documentos, o peso é calculado com base na distância entre o nome dos funcionários mencionados e os termos da consulta de busca, por exemplo, na frase “João apresentou um artigo no congresso Química e Sociedade” a distância entre o funcionário João e o termo química é 5(cinco). O resultado final, gerado a partir da combinação dos pesos das duas categorias de documentos, é exibido numa lista ordenada de especialistas que contém dados pessoais e documentos relacionados a cada funcionário, e é ordenada com base nos pesos calculados pelo sistema.

P@NOPTIC EXPERT (CRASWELL *et al.*, 2001) é um sistema “web” que identifica, de forma automática, os especialistas numa área através de seus documentos. Este sistema primeiramente monta o perfil de cada especialista através da concatenação do texto dos documentos associados, os quais estão publicados na rede interna. A seguir os perfis construídos são indexados e serão recuperados e ordenados conforme suas relevância na consulta de pesquisa. Somente os perfis dos 10 especialistas mais relevantes são retornados no resultado de busca. Enquanto o perfil do especialista mais relevante apresenta as informações de contato, foto e documentos relacionados ao especialista. Os perfis dos demais especialistas apresentam somente as informações de contato dos mesmos.

⁵ A Mitre Corporation é um centro independente de pesquisa e desenvolvimento global, sem fins lucrativos, que recebe verbas federais do governo americano e fornece engenharia de sistemas e conhecimento especializado em TI para órgãos públicos federais.

Em CNDS (YAO *et al.*, 2005), o objetivo é recuperar uma lista ranqueada de candidatos especialistas em um determinado tópico. Para atingir esse objetivo, o sistema emprega três métodos separadamente. O primeiro método corresponde à utilização de técnicas de Busca e Recuperação de Informação para criar consultas baseadas no conteúdo do tópico, recupera as páginas “web” mais importantes, extrai o nome das pessoas que estão nestas páginas, calcula a pontuação dessas pessoas e retorna uma lista ordenada das mesmas. O segundo método utiliza a técnica de clusterização de “e-mail” para analisar as principais características dos “e-mails” dos especialistas. Nesta técnica, os nomes dos remetentes são agrupados com base no relacionamento dos “e-mails” enviados, respondidos e recebidos. A seguir é criado um “link” entre o “cluster” e o tópico. Finalmente, uma lista de especialistas em cada tópico é gerada. O terceiro método extrai os textos âncoras e títulos de cada página e os utiliza para pegar as “entry pages” de cada tópico. Estas “entry pages” correspondem a tema principal de cada grupo de trabalho. Com base nas “entry pages” de cada grupo de trabalho e nas páginas pessoais, é criado o relacionamento entre pessoas e tópicos. A seguir é gerada a lista de especialista relevante em cada tópico. Por último, são utilizadas técnicas de agregação para combinar os resultados gerados pelos três métodos.

SAPIENS (SANTOS *et al.*, 2003) é uma ferramenta genérica, ou seja, independente de uma organização ou domínio específicos, e está baseada na infra-estrutura definida para Ambientes de Desenvolvimento de “Software” Orientados a Organização (ADSOrg⁶). O objetivo desta ferramenta é possibilitar que desenvolvedores de “software” encontrem rapidamente, dentro da estrutura organizacional, os profissionais mais adequados à realização de uma atividade ou à solução de um problema relativo ao desenvolvimento de um projeto de “software”. Esta ferramenta permite a representação da estrutura organizacional com as habilidades, conhecimentos e experiências requeridos ao longo dessa estrutura. Esta ferramenta também permite a alocação de pessoal, incluindo os conhecimentos, habilidades e experiências que cada profissional possui e disponibiliza

⁶ ADSOrg é definido como uma classe de ambientes de desenvolvimento de software que apóia a atividade de engenharia de software em uma organização, fornecendo o conhecimento acumulado pela organização e relevante para essa atividade, ao mesmo tempo em que apóia, a partir de projetos específicos, o aprendizado organizacional em engenharia de software (VILLELA *et al.*, 2001).

mecanismos para busca e navegação. O usuário pode realizar a busca através de uma série de consultas previamente cadastradas ou criar suas próprias consultas, as quais se baseiam nas relações existentes entre os conceitos que compõem a ontologia de organização. As informações referentes a cada usuário são cadastradas diretamente na ferramenta.

O NAVIGATOR (VIVACQUA *et al.*, 2006) é um sistema “peer-to-peer” que procura facilitar a colaboração oportunística em grandes redes sociais, em um mesmo contexto de trabalho. Este sistema constrói o perfil de atividades do usuário através da análise dos documentos existentes nos computadores dos mesmos. Um diferencial desse sistema é que ele considera o tempo de utilização dos documentos na construção e comparação dos perfis. O resultado da comparação é apresentado através de uma lista de possíveis contatos no computador de cada usuário. Os usuários só poderão interagir através dessa lista de contatos.

BECKS *et al.* (2004) desenvolveram um SRE para o “Fraunhofer e-Qualification framework” (E-QF, 2008), um ambiente de “e-learning” que oferece suporte técnico, metodológico e didático para autores e estudante de treinamentos via “web”. Este sistema constrói o perfil do especialista a partir de diferentes fontes de informação, tais como: histórico das interações (pessoas com quem se comunicou através de “e-mail”); percepção dos dados de seus documentos (materiais de aprendizado produzidos, editados e utilizados) e o próprio uso da plataforma de “e-learning” (buscas executadas, treinamentos que participa ou participou, tutorias realizadas etc), cada fonte de informação possui um algoritmo de comparação próprio. O sistema também possibilita a importação de materiais externos à plataforma que possam ser relevantes à construção do perfil. Neste sistema, o especialista decide quais dessas informações ele disponibilizará. Todo material da plataforma está categorizado numa Ontologia de conceitos, a qual é utilizada pelo algoritmo de comparação para retornar tutores e alunos relevantes aos materiais da categoria desejada. Este sistema é utilizado para dois propósitos: como *filtro de busca*, que permite um usuário encontrar outro que apresente perfil similar ao seu, ou como *funcionalidade de cluster*, neste caso o sistema agrupa os perfis de todos os usuários, para propósitos exploratórios, e os apresentam sobre a forma de “landscape of expertise”, conforme mostrado na Figura 28. Neste “landscape”, os usuários são representados por pontos, onde a proximidade dos pontos é decorrente da maior similaridade entre perfis.

Adicionalmente, variações de cinza são utilizadas para facilitar a identificação dos grupos formados, ou seja, pontos pertencentes ao mesmo grupo estão sobre a mesma tonalidade.

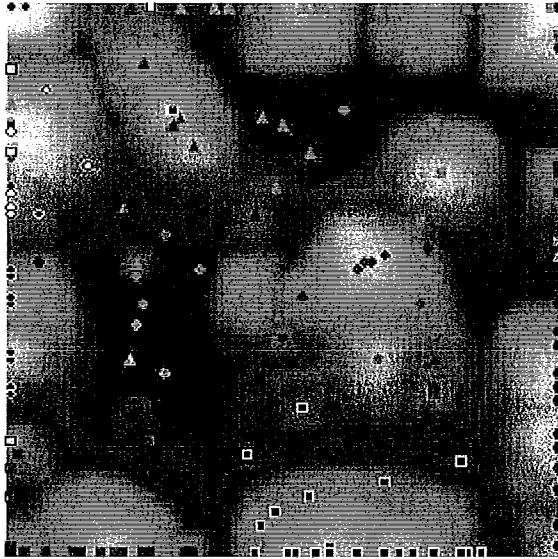


Figura 28 – “Expert landscape” para identificação de possíveis grupos (BECKS et al., 2004).

Dentre os SREs apresentados, somente o sistema SAPIENS não detecta dinamicamente as características utilizadas na comparação dos especialistas, pois estas são inseridas pelo próprio usuário. No caso dos demais sistemas, o NAVIGATOR é o único que não utiliza a página pessoal dos especialistas na detecção dinâmica de similaridades, os demais (EXPERT FINDER, E-QF, P@NOPTIC EXPERT e CNDS), além das páginas pessoais, utilizam outros documentos relacionados, tais como: relatórios, arquivos etc. CNDS e o E-QF também consideram as interações via e-mail nesse cálculo.

No que diz respeito à apresentação dos resultados retornados, SAPIENS, EXPERT FINDER, P@NOPTIC EXPERT e CNDS apresentam uma lista ordenada de especialista ao usuário. NAVIGATOR apresenta os resultados numa lista parecida com “msn”, a qual viabiliza a colaboração caso os especialistas recomendados estejam “online”. Finalmente, o sistema E-QF possibilita duas formas de retornos ao usuário. Se o sistema for utilizado como *filtro de busca*, os resultados são apresentados numa lista. Caso seja utilizado como *funcionalidade de cluster*, os resultados são exibidos na forma de um “landscape” de especialistas.

No que diz respeito à privacidade dos usuários. CNDS desconsidera totalmente este quesito, chegando a minerar informações diretamente da caixa de “e-mails” dos usuários, sem a percepção do mesmo. Os SAPIENS, EXPERT FINDER, P@NOPTIC EXPERT utilizam todas as informações publicadas, entretanto, depois de publicadas, essas informações não serão mais privadas. E-QF é o sistema que mais respeita a segurança e privacidade, permitindo que os usuários selecionem os itens que são analisados durante a busca de similaridades. Entretanto, essa abordagem tem um ponto negativo, pois quanto menor a quantidade de informação analisada, menor será a precisão das recomendações.

Por fim, NAVIGATOR é o único sistema que leva em consideração o tempo que um documento ficou aberto no cálculo de similaridades. Esse fator é um diferencial, pois representa um indício de que o usuário leu o documento.

Após esta análise, encerramos o estado da arte sobre os sistemas de recomendações. A seguir uma breve conclusão sobre o capítulo é descrita.

3.6. Conclusão

Este capítulo, com objetivo de contextualizar o usuário, apresentou o estado da arte de diferentes tipos de ferramentas de visualização da informação. Além disso, um estudo do estado da arte referente a sistemas de recomendação de especialistas também foi realizado, visto que esse é o objetivo da abordagem proposta nesta dissertação.

Devido a grande quantidade de informação envolvida na busca por um pesquisador substituto adequado, sentiu-se a necessidade de produzir uma ferramenta que possibilitasse uma maior agregação e facilidade de análise, da informação relevante ao processo de busca por substitutos. Logo, o estado da arte de visualização da informação possibilitou o surgimento de idéias referentes a que visualizações aplicar para chegarmos a estes objetivos. Enquanto o estudo da arte referente aos sistemas de recomendações permitiu à identificação dos pontos fortes e fracos dos sistemas selecionados, essas características foram muito úteis na concepção da ferramenta resultante desta dissertação.

O próximo capítulo descreve o trabalho realizado nesta dissertação, ou seja, uma abordagem que utiliza técnicas de visualização da informação para auxiliar a busca por pesquisadores no ambiente científico.

4. Busca de Especialistas através de Exemplo (BEE)

Essa dissertação propõe a abordagem BEE (*Busca de Especialistas por Exemplo*) cujo objetivo é facilitar a busca por pesquisadores substitutos em instituições de ensino e pesquisa, utilizando-se para isto técnicas de Visualização da Informação. Esta abordagem é composta de três passos (ter a informação na base de dados, descobrir contextos⁷ críticos do pesquisador a ser substituído e localizar os substitutos aptos) e foi implementada no ambiente GCC (Gestão do Conhecimento Científico). O primeiro passo é realizado utilizando-se as ferramentas existentes no próprio GCC, os outros dois são executados através do uso de técnicas de visualização.

Neste capítulo são apresentados a motivação e objetivos da abordagem proposta (seção 4.1), uma visão geral do ambiente GCC (seção 4.2), a abordagem proposta nesta dissertação (seção 4.3), além da comparação da implementação da BEE com os trabalhos correlatos apresentados no capítulo anterior (seção 4.4).

4.1. Introdução

A falta de recursos (tecnológicos, financeiros e de infra-estrutura), somada ao baixo número de concursos públicos, provoca grande mobilidade no meio científico. Conseqüentemente, os pesquisadores (alunos, mestres, doutores e professores concursados) permanecem pouco tempo em seus projetos de pesquisa, pois são incentivados a trocar de instituição de ensino ou largar a vida acadêmica para ingressar no mercado de trabalho. Desta forma, a instituição deve encontrar em seu próprio ambiente acadêmico um pesquisador apto a suprir essa perda.

Entretanto, encontrar um substituto adequado para um pesquisador não é uma tarefa trivial, pois envolve a análise de um grande volume de informação a respeito das partes: seus contextos de atuação, além de sua participação e produção (publicações bibliográficas, materiais, contribuições); competências; perfil; grau de conhecimento e capacidade de aprendizado. Essa sobrecarga faz com que os responsáveis gastem muito tempo analisando

⁷ Nessa dissertação, um contexto equivale a um projeto ou comunidade em que um pesquisador atua.

grandes volumes de dados ou baseiem sua escolha em recomendações de terceiros, fazendo com que na maioria das vezes o profissional escolhido não seja o mais indicado.

Portanto, uma solução que permita filtrar, processar e organizar toda essa informação trará um grande benefício aos responsáveis por encontrar tais substitutos. Afinal, quanto mais concisa e organizada for a informação analisada, maior será a probabilidade de se escolher o substituto correto. Na abordagem desenvolvida nesta dissertação, a informação envolvida no processo de escolha do substituto foi classificada nos três espaços propostos por JUDELMAN (2004):

1. Espaço de informação complexo: contém as informações referentes aos contextos de atuação do pesquisador, devido à natureza hierárquica das mesmas.
2. Espaço de informação de contexto: compreende as competências e perfil MBTI (MYERS, 1980), pois estes representam as similaridades de conhecimento e psicológicas dos pesquisadores.
3. Espaço de informação dinâmica: é formado pelo grau de conhecimento e a capacidade de aprendizado do pesquisador, já que estes fatores são analisados temporalmente.

Esta dissertação descreve a abordagem BEE (Busca de Especialistas através de Exemplo). O objetivo dessa abordagem é apoiar a substituição de pesquisadores, ou seja, identificar pessoas aptas a realizarem as mesmas tarefas do pesquisador a ser substituído, utilizando-se para isto técnicas de visualização. Para isso, a BEE identifica os contextos que sofrerão maior impacto após a saída do pesquisador e recomenda, a partir destes, substitutos conforme a similaridade entre suas competências e as do profissional a ser substituído.

A BEE é sensível ao contexto do pesquisador a ser substituído. Nesta dissertação, consideramos como contexto as comunidades, projetos e a instituição de um pesquisador. Nossa abordagem usa o conhecimento contextual (as competências relevantes ao contexto da busca) para filtrar os possíveis substitutos, ou seja, somente pesquisadores que possuem alguma competência, relevante ao contexto e comum a do profissional a ser substituído, são considerados. Os candidatos substitutos são classificados e apresentados em função do grau de similaridade com o substituído. Caso o usuário permaneça na dúvida de qual substituto

escolher, a BEE disponibiliza outros critérios de seleção, tais como: grau de interação com um tema de interesse, nível de aprendizado numa competência e comparação do perfil MBTI.

A BEE propõe o uso de técnicas de “Visualização de Informação” pertencentes aos três espaços de informação propostos por JUDELMAN (2004) - complexo, contexto e dinâmico conforme mostrado nas seções 3.2, 3.3, e 3.4, respectivamente - para apoiar o processo de substituição anteriormente descrito. Essas técnicas proporcionam uma representação clara e concisa dos valores dos parâmetros que o usuário, apoiado na visualização da atuação do pesquisador em diversos contextos, considerou mais relevantes para comparação entre os substitutos candidatos.

Para avaliar esta abordagem foi implementada uma ferramenta, denominada FBEE (*ferramenta para busca de especialista através de exemplo*), que foi acoplada como um novo módulo ao GCC. Não é objetivo da ferramenta realizar a coleta dos dados a serem analisados, essa tarefa é realizada através do GCC. Apesar deste ambiente possuir toda informação necessária à busca de um pesquisador substituto (competências, áreas de interesse, publicações, projetos, comunidades etc), a forma de execução dessa busca não é eficaz, pois toda a análise para se escolher um profissional é baseada em relatórios. Além disto, muitas informações relevantes são desconsideradas durante a análise relatorial dos possíveis substitutos, como por exemplo, a importância do pesquisador na instituição de ensino. Como o GCC não disponibiliza ferramentas adequadas para a análise dos três espaços de informação, um grande esforço cognitivo é requerido para encontrar os substitutos adequados.

O conhecimento contextual, apoiado por técnicas de *Visualização de Informação*, permite que a FBEE faça recomendações de candidatos substitutos de maneira mais coerente, tendo em vista que o usuário terá uma visão clara e concisa de toda informação relevante existente no GCC, além de poder controlar os dados exibidos e organizá-los de modo a facilitar seu entendimento. A próxima seção apresenta o ambiente em que a abordagem BEE foi implementada.

4.2. Ambiente GCC

O GCC é um ambiente “web”, desenvolvido inicialmente para atender ao público interno da COPPE e posteriormente expandindo para atuar em todo cenário científico e acadêmico brasileiro. O GCC propicia uma centralização de informações do meio acadêmico, facilitando a disseminação do conhecimento gerado. Além disto, este ambiente disponibiliza toda a infra-estrutura para a criação e manutenção de comunidades virtuais de pesquisa e gerência de projetos, estimulando assim o desenvolvimento de novas idéias e novos trabalhos e a colaboração entre os pesquisadores (OLIVEIRA *et al.*, 2005). A Figura 29 mostra uma visão geral da arquitetura do GCC e todos os seus módulos, os quais são detalhados nas seções a seguir.

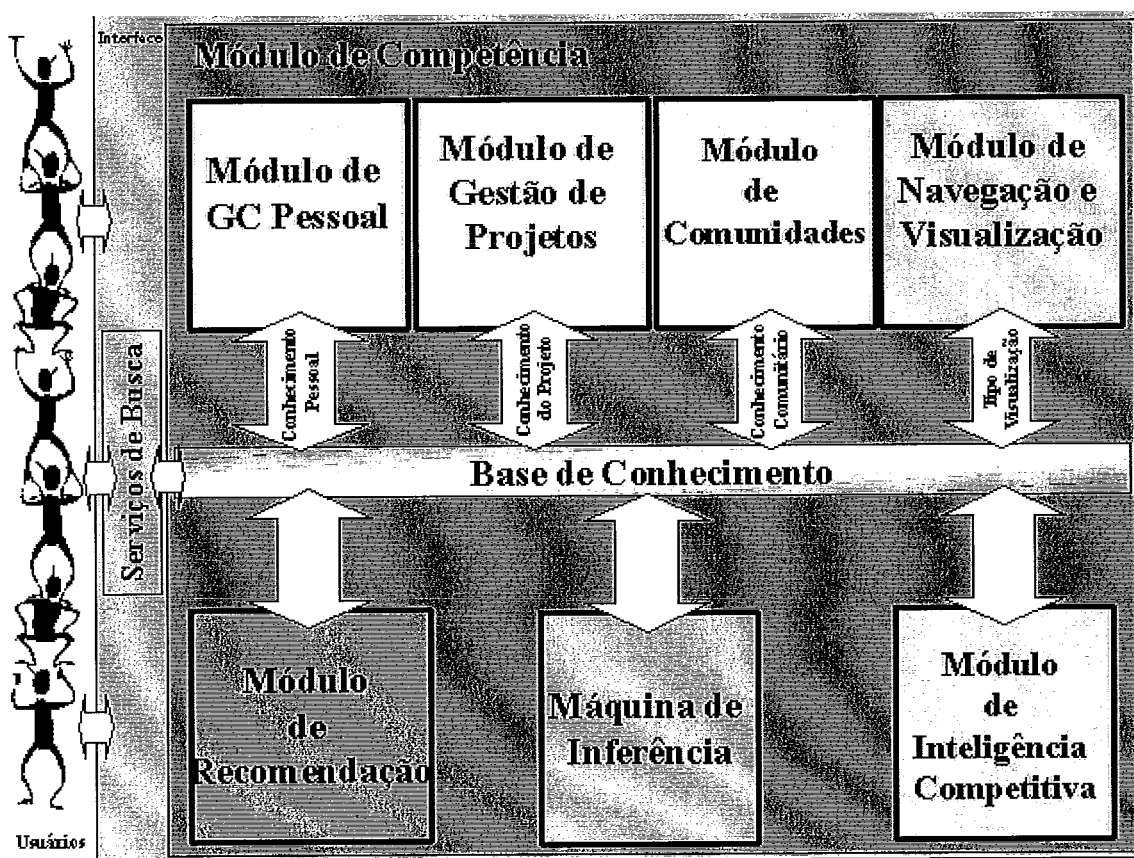


Figura 29 – Visão Geral da Arquitetura do GCC.

4.2.1. Módulo de Gestão de Competências

Um grande problema enfrentado na Gestão do Conhecimento é a incapacidade da organização em descobrir o que ela sabe, ou seja, as competências que domina. Esta descoberta pode auxiliar bastante o processo de tomada de decisão, a formação de grupos de trabalhos e o aprimoramento organizacional. Este problema é aumentado no ambiente científico pela liberdade na escolha de temas de pesquisa, pouca hierarquia, falta de estratégia e correlação dos trabalhos produzidos em uma instituição científica.

Atualmente, no GCC, a competência pode ser declarada pelo pesquisador, inferida pelo uso do GCC, extraída do currículo Lattes ou ainda mapeada pela mineração dos textos publicados pelos pesquisadores. Para isto, este módulo conta com outros sub-módulos, que são descritos a seguir.

4.2.1.1.S-Miner: Mineração de Competências

Uma abordagem para a descoberta de competências de um pesquisador é utilizar suas publicações. Após identificar os assuntos relevantes de uma publicação, pode-se inferir quais destes são competências dos autores desta publicação, tendo em vista que escreveram sobre o assunto. Desta forma, no GCC, os textos publicados são submetidos a um minerador e, ao final do processo de extração, é retornada uma lista de possíveis competências que o pesquisador domina, como mostrado na Figura 30. Estas competências são dispostas junto às demais (já contempladas no GCC) e armazenadas na base de dados para futuras pesquisas. O minerador utilizado nesta aplicação é denominado S-MINER e foi desenvolvido por RODRIGUES *et al.* (2004). Atualmente, esta ferramenta também está sendo utilizada na mineração de competência dos Currículos Lattes dos pesquisadores.

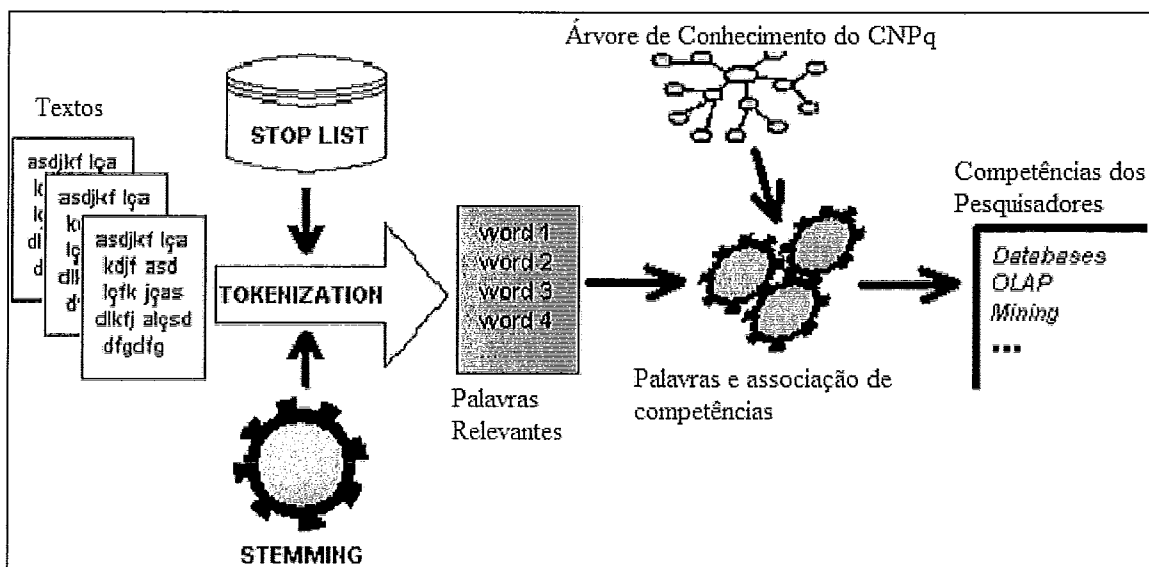


Figura 30 – Arquitetura do S-MINER (RODRIGUES et al., 2004).

Ao final de todo o processo de mineração, as competências extraídas são vinculadas aos respectivos autores (pesquisadores) e disponibilizada nos diversos relatórios que envolvem competência do GCC, especialmente na Busca de Competências e na Busca de Publicações, abordados a seguir.

4.2.1.2. Busca de Competências

A arquitetura da busca é baseada nos dados do GCC, incluindo as competências mineradas e nos dados encontrados no currículo Lattes. As competências são pesquisadas conforme os seguintes critérios:

- Competências Declaradas – são as competências que o próprio pesquisador julga ter e no seu suposto grau (alto, médio ou baixo) de “expertise”. Estas competências são cadastradas no Módulo do Pesquisador no GCC pelo próprio usuário.
- Competências de Projetos – são as competências encontradas nos projetos dos quais o pesquisador faz parte ou já participou. Estas competências são armazenadas no momento de criação do projeto, no Módulo Projeto do GCC.
- Competências Extraídas – são as competências identificadas pelo S-Miner, o qual foi descrito na anteriormente.

- Competências de Comunidades – são as competências definidas pela(s) comunidade(s) a que o pesquisador pertence, as quais são declaradas no Módulo de Comunidade do GCC.

Alguns critérios para identificação de competências, conforme mostrado na Figura 31, são extraídos do currículo Lattes a partir dos títulos e descrições da produção bibliográfica, produção técnica, orientações e demais informações que são minerados em cada currículo.

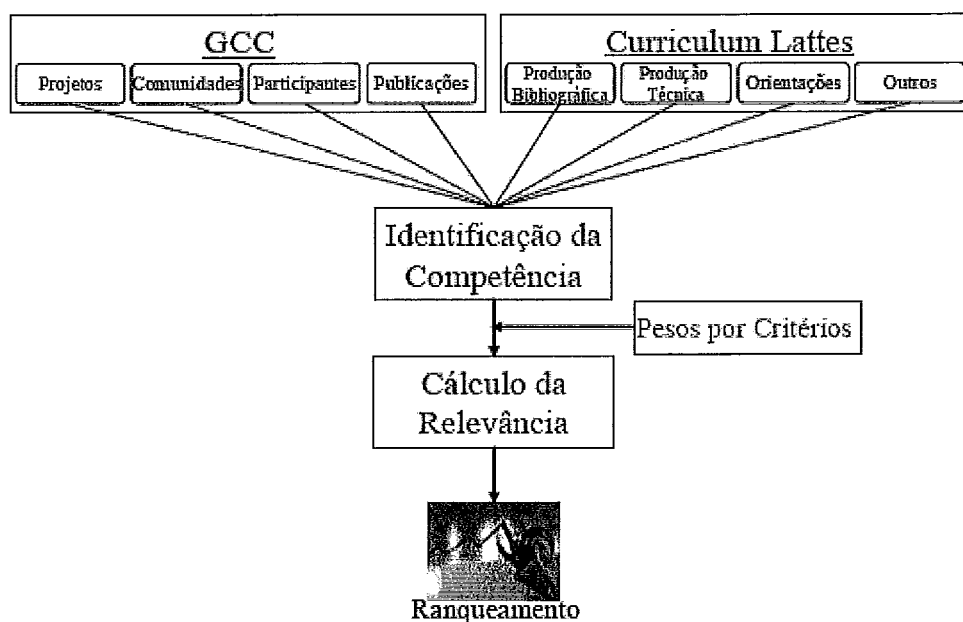


Figura 31 – Fases da Busca da Competência (OLIVEIRA, 2007).

Aliado a estas formas de mapear as competências, este sub-módulo de Busca de Competências considera pesos diferenciados para cada tipo de competência encontrada, os quais são definidos livremente pelo usuário. Ao contabilizar essas competências, este sub-módulo identifica na Árvore do Conhecimento do CNPq os níveis em que cada tipo de competência se encontra. Ao descobrir estes níveis, é possível mapear a relevância de cada pesquisador mediante as competências buscadas e, desta forma, reportar uma lista com os pesquisadores que possuem as habilidades procuradas. Na Figura 32, encontra-se um exemplo de busca de competências e seus resultados.

Gestão de Conhecimento Científico

Adminstrador: J. Jones de Oliveira de Aguiar

Projeto | Participante | Comunidade | Visualização | Competência | Relevância Competitiva | Ferramentas Administrativas | Sair

Busca por Competências

Competências Disponíveis (Área de Conhecimento)

- Eng. Banco de Dados
- Eng. Respostas (Ergonomista da Sobrep)
- Geometria Construída (Física Matemática)
- Gestão do Conhecimento & Intuições de Ensino (Banco de Dados)
- Gestão do Conhecimento (Banco de Dados)
- Gnd (Banco de Dados)

Competência escolhida

Critérios de Seleção

Gestão do Conhecimento (Banco de Dados)

Áreas de Seleção

Condições: Gestão do Conhecimento (Banco de Dados)

Relevância: 1

Projeto: 1

Publicações: 1

Experiência: 1

Serviços Bibliográficos: 1

Produção Técnica: 1

Cititações: 1

Outros: 1

Peso dos item

Critérios utilizados no cálculo da relevância

Pesquisadores que possuem a competência.

Nome	Relevância	Experiência	Projeto	Publicações	Serviços Bibliográficos	Produção Técnica	Cititações	Outros
Melissa de Sousa	2	✓	✓					
Marcelo	4	✓	✓	✓	✓			
Janaína de Sousa	4	✓	✓		✓			
Marcelo de Sousa	4	✓	✓		✓			
Carla de Sousa	2	✓						
Carla de Sousa	2	✓						
Carla de Sousa	2	✓						
Carla de Sousa	2	✓						
Carla de Sousa	2	✓						
Carla de Sousa	2	✓						
Carla de Sousa	2	✓						
Carla de Sousa	1							
Carla de Sousa	1							

Grau de conhecimento na competência

Figura 32 – Busca de Competências (OLIVEIRA, 2007).

A seguir é apresentada uma outra busca que faz uso das competências mineradas pelo S-MINER.

4.2.1.3. Busca de Publicações

O resultado da *busca de publicações* é baseada nos textos minerados previamente pelo S-Miner. A partir de uma ‘interface’ de consulta, é possível indicar quais competências que deverão ser procuradas (Critérios de Seleção). O resultado é apresentado em ordem de relevância. O procedimento para obtenção desta relevância é feito a partir da classificação de competências (designada aos técnicos especializados) e ao mapeamento de palavras relevantes extraídas dos textos.

Esta seção encerra o módulo de competências, na próxima seção é descrito o módulo de *Gestão de Conhecimento Pessoal*.

4.2.2. Módulo de Gestão de Conhecimento Pessoal

Pesquisadores costumam realizar as suas anotações em cadernos, bloco de folhas ou até mesmo em folhas avulsas. Para tentar reproduzir esta realidade, o módulo de *Gestão do Conhecimento Pessoal* foi construído.

Este módulo provê ferramentas para um pesquisador gerenciar o seu conhecimento pessoal e as informações a respeito dele próprio. A seguir são listadas as principais ferramentas disponibilizadas por este módulo:

1. “Curriculum Vitae”: É uma das formas de manter informações sobre os profissionais da organização.
2. Diários “Web”: são as páginas pessoais dos pesquisadores no GCC. O *diário web* é mais uma ferramenta para propiciar a gestão do conhecimento pessoal.
3. Mapas Mentais: representam como uma pessoa vê o mundo a sua volta. No GCC, os pesquisadores podem utilizar essa ferramenta para facilitar a expressão de idéias, auxiliar o “brainstorm” e como subsídio à preparação de reuniões.

A partir do momento que informações passam a ser registrada em sistemas e não mais em papel, a chance que elas se percam diminuem consideravelmente, além da organização que passaram a ter.

Conhecimentos provenientes de projetos também devem ser gerenciados, o módulo a seguir foi desenvolvido com o objetivo de apoiar essa questão.

4.2.3. Módulo de Gestão de Projetos

Tendo em vista que organizações de pesquisa possuem e desenvolvem projetos (sendo que em cada projeto pode-se ter vários processos relacionados), cada projeto, para o seu bom desempenho, necessita de conhecimento “certo” no momento “certo”. Adicionalmente, muitos conhecimentos são gerados na criação e execução de projetos. Logo, os projetos são importantes fontes de conhecimento, como também são mecanismos de aplicação de conhecimento.

Esse módulo provê ferramentas para auxiliar a criação e execução de projetos. Os responsáveis pela coordenação de um projeto criam o modelo de processo, com a seqüência

de atividades e os insumos necessários para cada uma destas atividades: áreas de conhecimento do CNPq, competências, documentos e outros artefatos. O modelo do processo é criado através de uma ferramenta gráfica de “workflow” e a sua execução é controlada por uma máquina de “workflow” (BOMFIM, 2005). Adicionalmente, esse módulo possibilita o armazenamento e distribuição do conhecimento gerado em um projeto. Como conhecimento de um projeto, entende-se:

1. Modelos de Processos: a seqüência de atividades é um conhecimento que pode ser reusado. Aproveitando-se “workflows” ou modelos de processos bem sucedidos economiza-se tempo, evitando-se tentativas de soluções não testadas e comprovadas, reduz a quantidade de erros, bem como privilegia a economia de recursos.
2. Definição de Atividades e Tarefas: atividades e tarefas bem definidas resultam em atividades e tarefas bem executadas.
3. O conhecimento adquirido ao término de uma tarefa e/ou atividade: um erro, uma nova prática, material adicional, idéias, sugestões podem ser cadastrados, no GCC, pelo usuário ao final da execução de uma tarefa ou ao término de uma atividade.

Comunidades, assim como projetos, são fontes geradoras de conhecimento. O próximo módulo do GCC foi desenvolvido com o intuito de facilitar o gerenciamento dos conhecimentos oriundos de comunidades.

4.2.4. Módulo de Comunidades

Um dos pontos principais do GCC é a criação de comunidades virtuais para que as pessoas discutam, aprendam e possam trabalhar juntas. No GCC, comunidades são grupos de pesquisadores com interesses comuns que podem trocar informações e trabalhar de forma colaborativa. O módulo de comunidades provê ferramentas de colaboração assíncrona e síncrona (ferramentas de reunião eletrônica, enquête, biblioteca digital, troca de mensagens, fórum e disponibilização de eventos e notícias), o que facilita a disseminação de informação e conhecimento entre membros de uma comunidade. Todas as informações e conversas realizadas através deste módulo são armazenadas, já que

representam uma forma de explicitação do conhecimento que pode ser consultada e analisada no futuro.

Este módulo ainda possui mecanismos de busca, permitindo ao usuário pesquisar comunidades por palavras-chave encontradas no nome e descrição da comunidade, seus participantes, áreas de conhecimento e competências que são assunto da comunidade e período de criação.

O módulo apresentado a seguir foi desenvolvido para facilitar o uso do ambiente GCC.

4.2.5. Módulo de Navegação e Visualização

A visualização típica na maioria dos sistemas propõe uma estrutura pobre de interfaces amigáveis, ausência de estruturas intuitivas de consultas e eventualmente existe a necessidade do usuário aprender linguagens ou convenções para interagir efetivamente com os sistemas. Baseando-se nestes problemas, além dos relatórios, listas e visualização gráfica provida no GCC, o ambiente provê três ferramentas adicionais para navegação e visualização:

1. Mapa Conceitual do projeto: apresenta o conhecimento e/ou informações e seus relacionamentos de uma maneira visual e mais intuitiva (diferentemente de relatórios comuns), permitindo que o usuário interaja, navegue e acesse as informações e conhecimentos.
2. Fluxo de Conhecimento: representa como um conhecido é gerado a partir do outro, a relevância de um conhecimento em um contexto e como os diferentes conhecidos se relacionam.
3. Árvore Hiperbólica: É uma interface de visualização de uma estrutura hierárquica em árvore. Na árvore hiperbólica do GCC, é possível visualizar as áreas de conhecimento de acordo com o CNPq, conforme mostrado na Figura 33. Um usuário pode navegar pela classificação e consultar os projetos, competências e outras informações associadas com a área de conhecimento selecionada.

4.2.7. Módulo de Inferência

Por meio de estratégias específicas, informações ocultas e previamente desconhecidas podem ser extraídas de grandes bases de dados na forma de regras e padrões.

O objetivo deste módulo é descobrir regras de associações entre os dados do GCC, tais como: Competência X Pesquisador X Tempo; Competência X Comunidade; Competência X Comunidade X Tempo; Comunidade X Pesquisador X Tempo; Competência X Tempo; Competência X Publicação; Publicação X Pesquisador; Publicação X Tempo; Projeto X Competências; Projeto X Pesquisador; Pesquisador X Pesquisador, que possam dar apoio à tomada de decisões.

O próximo módulo é responsável pela análise comparativa e posicionamento de uma pessoa, departamento ou instituição no seu ambiente.

4.2.8. Módulo de Inteligência Competitiva

Uma análise do *mercado acadêmico* tende a apontar os pontos fortes e fracos da instituição quando se analisa o ambiente interno, sendo possível verificar as ameaças ou oportunidades quando se investiga o ambiente externo. A partir dessas análises, a instituição consegue se posicionar de uma melhor forma, utilizando para isto situações identificadas no ambiente interno ou externo.

O módulo de IC tem como objetivo geral permitir que o pesquisador, linha de pesquisa, departamento ou instituição conheça melhor a si mesmo e aos seus *competidores*. O ambiente que envolve a instituição é monitorado com o intuito de aumentar a vantagem competitiva na mesma. Através dos indicadores, são feitas análises comparativas por pesquisador, por setor (linha de pesquisa ou departamento) e por instituição, tornando-se possível detectar pontos fracos e pontos fortes, o que permite a elaboração de ações que acarretem em melhorias significativas para a instituição (KAWAMURA, 2006).

Com as informações do módulo de IC também é possível conhecer padrões da atividade em C&T, tais como projetos de pesquisa, produções, priorização de temas de trabalho, entre vários outros itens. O que torna possível traçar um panorama da pesquisa científica dentro do escopo analisado.

A descrição deste módulo finaliza a seção referente ao GCC, na próxima seção apresentaremos a descrição da proposta deste trabalho.

4.3. Descrição da Proposta

A localização de pessoas aptas a substituir um determinado especialista, na execução de uma determinada tarefa ou na ocupação de um cargo, é um problema frequente nas grandes organizações (YUPENG *et al.*, 2006). No cenário científico, a saída de uma pessoa pode ser decorrente de vários fatores, como a finalização do curso, obtenção de um estágio ou emprego, um doutorado-sanduíche, pós-doutorado, aposentadoria ou término da bolsa. Isto resulta em equipes altamente inconstantes.

Com isto, para evitar maiores danos na execução de um projeto de pesquisa é necessário que a alocação de membros seja feita de maneira otimizada. Ou seja, a pessoa que ocupará o novo lugar no projeto deve ser o mais similar possível com o membro que saiu (ou sairá), ou ao menos, se encaixar com as exigências requeridas.

Com objetivo de solucionar esse problema, esta dissertação criou a abordagem BEE, a qual teve uma ferramenta implementada para que pudesse ser utilizada e avaliada. A implementação tornou-se o sub-módulo Busca de Substituto no GCC e foi denominada FBEE (ferramenta de busca de especialista através de exemplo).

A BEE baseia-se numa busca composta de três fases: ter a informação na base de dados, descobrir contextos críticos e busca por substitutos.

1. Ter a Informação na Base de Dados: nesta fase, toda informação referente ao pesquisador é armazenada e associada às suas competências, através do GCC. Por exemplo, as comunidades, projetos, atividades ou publicações do pesquisador são relacionadas a uma ou mais competências do mesmo. Esse armazenamento/associação pode ser realizado de forma automática (durante a importação do Lattes), semi-automática (através da mineração de competências via S-Miner) e de forma manual (pelo próprio usuário do sistema). Apesar dessa fase ser realizada externamente a BEE, ela foi acrescentada a abordagem porque não seria possível aplicá-la sem executar esta fase.

2. Descobrir contextos críticos: nesta fase, o usuário define o contexto de atuação do pesquisador que sofrerá maior impacto após sua saída.
3. Busca por substitutos: nesta fase, ocorre a indicação dos possíveis pesquisadores substitutos. Essa indicação se baseia na similaridade do grau das competências, no perfil MBTI, na capacidade de aprendizado e no nível de interesse dos mesmos.

A abordagem BEE emprega técnicas de Visualização de Informação para apoiar essa abordagem de busca. Essas técnicas englobam os três espaços de informação: complexo, contexto e dinâmica propostos por JUDELMAN (2004), conforme descrito na seção 3.1. Estes três espaços de informação geram uma maior precisão nos resultados das buscas executadas. Logo, esta abordagem permite ao usuário:

1. Visualizar os projetos, comunidades e a instituição nos quais um determinado pesquisador participa e a importância dele nos mesmos (espaço de informação complexo). Com isso, o usuário consegue identificar quais destes sofreram maior impacto após a saída do pesquisador, podendo assim, priorizar a substituição nos mesmos. Essa visualização é feita através de *Mapa de Árvores*, vide seção 4.3.2 e está incluída na segunda etapa da abordagem.
2. Comparar, através de visualizações, as características do pesquisador a ser substituído com as dos seus substitutos (espaço de informação de contexto). Essa comparação é realizada através do *gráfico de teia ou área* e do *gráfico de força*, vide seções 4.3.3.1 e 4.3.3.2 respectivamente, e está incluída na terceira etapa da abordagem. O uso de visualizações reduz o esforço cognitivo necessário a essa análise, pois as informações, além de já estarem processadas, são apresentadas de modo intuitivo ao avaliador.
3. Comparar o fluxo de aprendizado dos pesquisadores (espaço de informação dinâmico). Essa comparação permite ao avaliador identificar quais pessoas estão trabalhando mais recentemente em um assunto, permitindo assim que as escolhas sejam feitas com base nos interesses atuais. Essa comparação é realizada através de *gráficos de linha*, os quais variaram conforme o período de tempo estipulado pelo usuário, vide seção 4.3.3.2, e também está incluída na terceira etapa da abordagem.

O “workflow” da Figura 34 ilustra as etapas seguida pela abordagem BEE.

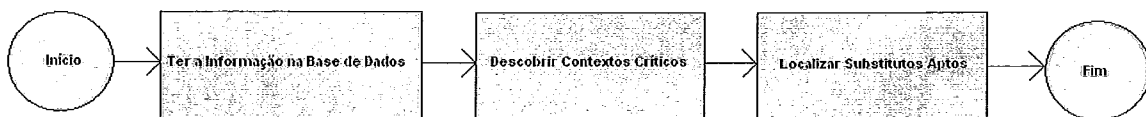


Figura 34 – “Workflow” das etapas da abordagem BEE.

Os componentes desse “workflow” são descritos nas subseções a seguir.

4.3.1. Ter a Informação na Base de Dados

Para o correto funcionamento da ferramenta, algumas informações referentes a pesquisadores, projetos e comunidades precisam ser especificadas no GCC:

1. Informações de Pesquisadores:

- Interesses – Conjunto de temas, assuntos, de importância para o usuário, mas nos quais ele ainda não possui conhecimento. Esta informação é provida pelo próprio usuário. Para cada interesse o usuário deve mencionar o grau do seu interesse no assunto. O grau varia em alto, médio ou baixo.
- Áreas de Conhecimento e Competências – Áreas da Ciência, definidas pelo CNPq, e competências nas quais o usuário possui algum conhecimento. O grau de conhecimento em uma Área do CNPq ou em uma competência varia em alto, médio ou baixo. O próprio usuário pode informar estas informações, sendo que o ambiente tenta inferir competências que o usuário tem, conforme descrito no Módulo de Competência.
- Perfil MBTI – Para a identificação da personalidade é utilizado o teste MBTI. Este teste é executado através das respostas dadas pelo usuário ao questionário existente no GCC.

Essas informações são gerenciadas no módulo de Gestão do Conhecimento Pessoal do GCC, conforme mostrado na seção 4.2.2.

2. Informações de Projetos

- Áreas de Conhecimento e Competências – Áreas da Ciência, definidas pelo CNPq, e competências que estão envolvidas ou são temas de um projeto. No

projeto, o grau de conhecimento em uma Área do CNPq ou em uma competência varia de 1 a 10. O próprio usuário, no caso gerente do projeto, insere essas informações.

- Tarefas – Correspondem às atividades desempenhadas pelos pesquisadores em um projeto. Para cada tarefa existem os pré-requisitos, ou seja, áreas de conhecimento ou competências necessárias para a realização da tarefa com sucesso. Tais áreas ou competências recebem um grau que varia de 1 a 10. Essas informações são cadastradas pelo usuário no sistema, no caso, o gerente do projeto.
- Materiais, Publicações e Contribuições – correspondem aos artefatos (tais como: artigos, relatórios, textos) criados ou consumidos pelos membros do projeto. Cada um desses artefatos é associado a uma ou mais competências, as quais recebem um grau de 1 a 10.
- Participantes e Cargos – todo projeto possui pessoas, as quais possuem cargos. Logo, essas informações também se fazem necessárias ao GCC. Diferentemente das outras informações, não existe relacionamento direto entre o cargo da pessoa e suas respectivas competências.

Essas informações são gerenciadas no módulo de Gestão de Projetos, conforme mostrado na seção 4.2.3.

3. Informações de Comunidades:

- Áreas de Conhecimento e Competências – Áreas da Ciência, definidas pelo CNPq, e competências que estão relacionadas à comunidade. Em uma comunidade, o grau de conhecimento em uma Área do CNPq ou em uma competência varia de 1 a 10. Essas informações são inseridas pelo moderador da comunidade.
- Participantes e Função – No GCC, um membro de uma comunidade pode ser moderador ou apenas um integrante da mesma.
- Contribuições: correspondem aos materiais, enquetes, eventos, “links”, notícias, mensagens em fóruns que um participante disponibilizou em sua comunidade. Nesse caso, somente materiais e fóruns são associados a competências, as quais

recebem um grau entre 1 e 10. Todas essas contribuições são inseridas pelos membros de uma comunidade.

Essas informações são gerenciadas no módulo de Gestão de Comunidades, conforme mostrado na seção 4.2.4.

Apesar da etapa *Ter a Informação na Base de Dados* não fazer parte da abordagem proposta, sem esta não seria possível executar a ferramenta de BEE, dado que esta não permite a inserção de dados. Por esse motivo, essa etapa foi inserida no “workflow” da abordagem. Esta etapa representa a transformação dos “Dados Brutos” em “Tabelas de Dados” do Modelo de Referência para Visualização explicado na seção 2.2. A seguir descrevemos o segundo passo da abordagem BEE.

4.3.2. Descobrir Contextos Críticos

Após o preenchimento da base, pode-se iniciar o processo da busca de substituto. O objetivo da etapa *descobrir contextos críticos* é facilitar a detecção dos contextos que poderão ser mais afetados após a saída do pesquisador.

O primeiro passo é selecionar a pessoa que desejamos substituir. A partir daí é gerada uma visualização contendo todos os contextos de atuação dessa pessoa. Existem três possíveis contextos de busca para o usuário. O contexto de projeto é utilizado quando o usuário deseja encontrar um substituto dentro do próprio projeto que o pesquisador abandonou. Já o de comunidade é utilizado para encontrar um substituto dentro da própria comunidade. Por fim, o contexto organizacional é utilizado quando se deseja encontrar o melhor substituto, independente dele participar ou não dos projetos e comunidades da pessoa a ser substituída.

A visualização escolhida para essa etapa foi o *Mapa de Árvores*, já descrito na seção 3.2.1.2. *Mapa de Árvores* são mapas criados para a visualização de uma grande quantidade de dados, apresentando as informações baseadas no tamanho e cor dos retângulos de acordo com algum atributo. Este tipo de visualização é utilizado no GCC para demonstrar a importância de uma pessoa em um determinado contexto. A cor do retângulo representa a relevância do pesquisador no seu contexto de atuação, ou seja, quanto mais claro for a cor na escala das cores, maior é a relevância do pesquisador neste contexto. O tamanho do retângulo indica o grau de participação, isto é, retângulos de maior área

representam profissionais mais participativos. Logo, quanto maior for a área e maior for a cor na escala das cores, maior é a importância do pesquisador no contexto escolhido. Portanto, essa visualização possibilita ao usuário descobrir, de uma maneira rápida e direta, quais projetos e comunidades sofrerão maior impacto após a saída do pesquisador, além da sua real importância na instituição. Além disso, esta fase funciona como um filtro de informação, permitindo que na etapa seguinte sejam analisadas somente as competências relevantes ao contexto selecionado.

O primeiro nível da visualização mostra as relevâncias das pessoas em relação a todas as comunidades e projetos existentes, além da importância de cada pessoa no contexto geral, em comparação aos demais usuários do sistema. Partindo-se desta visualização, pode-se especializar a consulta navegando-se pelos projetos ou comunidades e ver a importância das pessoas nos mesmos ou em relação aos demais usuários do sistema, conforme ilustrada na Figura 35.

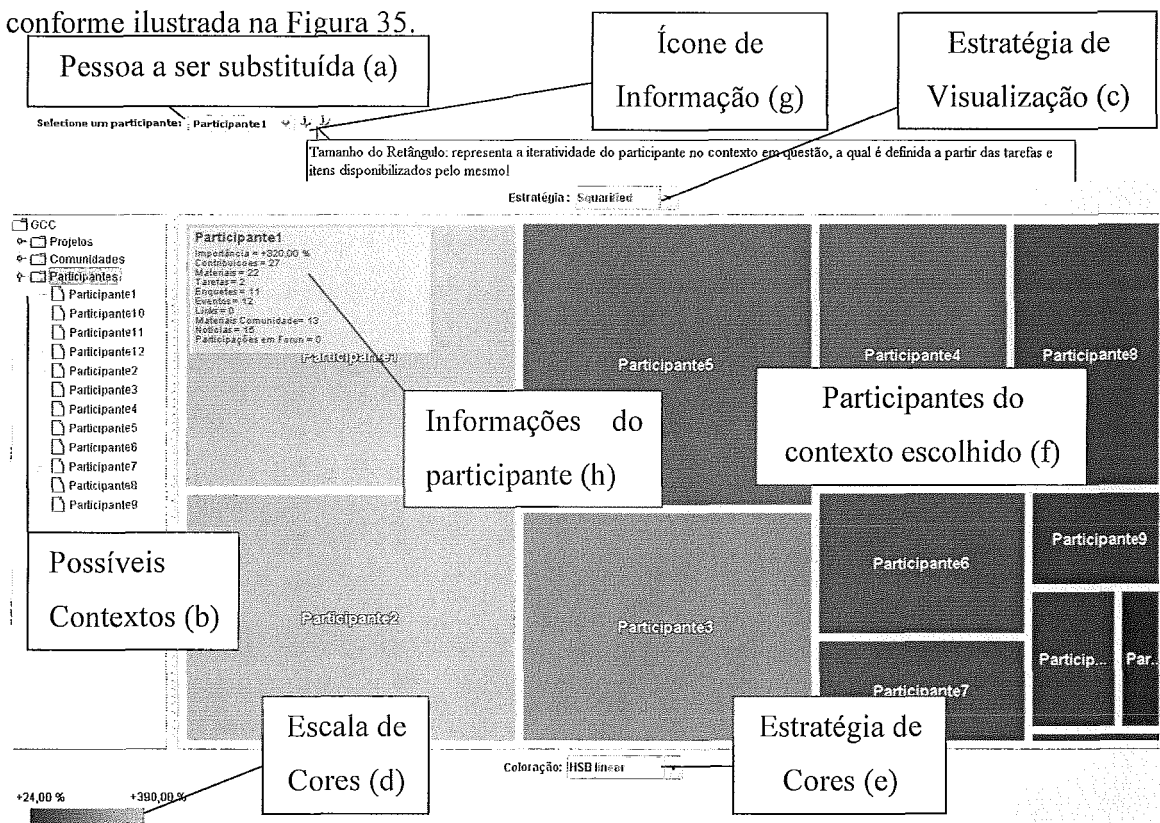


Figura 35 – Relevância das pessoas na organização.

Maiores detalhes da visualização estão na Figura 35. Para este exemplo, foi escolhida a pessoa para quem queremos encontrar um substituto (Figura 35, item a) no

contexto organizacional (Figura 35, item b). O algoritmo de construção do Mapa de Árvores e a escala de cores dos retângulo podem ser alterados conforme o gosto do usuário (Figura 35, item c e d, respectivamente). A parte central da visualização (Figura 35, item f) mostra as pessoas mais relevantes do contexto escolhido. Os ícones de informação (Figura 35, item g) informam o significado da cor e tamanho do retângulo. Por último, quando se posiciona o cursor do “mouse” sobre um dos retângulos, uma listagem contendo as contribuições e importância do participante no contexto em questão é apresentada (Figura 35, item h).

O cálculo da cor (representado pelas funções PR, CR e UR) e área do retângulo (representado pelas funções PI, CI e UI) em cada contexto do *Mapa de Árvores* é detalhado a seguir:

1. Contexto Projeto:

a. Conjunto de informações utilizadas:

- Função do participante no projeto – Caso o participante seja gerente no projeto considerado, a sua importância será aumentada.
- Número de tarefas destinadas ao participante – Quanto maior for o número de tarefas destinadas ao participante, maior a sua importância.
- Competências – Competências e grau destas competências que o participante possui e que são equivalentes às competências relevantes ao projeto. Neste caso, é feito um comparativo do grau de cada competência exigida pelo projeto com o grau das competências pessoais de cada participante. Assim, quanto maior for a proximidade do grau destas competências, maior será a importância da pessoa para o projeto.
- Contribuição – Mede a interatividade e apoio do participante para os demais membros do projeto através da submissão de idéias, críticas, sugestões, notícias e disponibilização de materiais úteis para o projeto. Quanto maior for a contribuição da pessoa para o projeto, maior é a sua importância.

- b. A cor do retângulo é determinada pela Equação 1. A função $PR(x, y)$ representa a relevância do pesquisador “x” no projeto “y”. $CompPesq(x, y)$ representa o valor da competência do pesquisador “x” que é similar a uma competência do projeto “y”, enquanto $CompProj(y, x)$ representa o valor da competência do projeto “y” que é similar a competência do pesquisador “x”. $Func(x, y)$ representa a função do pesquisador “x” no projeto “y”. $Func(x, y)$ será 50 caso o pesquisador seja o gerente do projeto e 0 nos outros casos, o valor 50 foi escolhido, pois indica uma variação perceptível na cor da escala das cores. Vale lembrar que $CompPesq(x, y)$ varia entre 1 e 3 (baixo, médio e alto) e $CompProj(y, x)$ entre 1 e 10, por isso consideramos duas notações distintas.

Equação 1 – Relevância (cor) do pesquisador em um projeto.

$$PR(x, y) = (\sum CompPesq(x, y) * CompProj(y, x)) + Func(x, y)$$

- c. O tamanho do retângulo é determinado pela Equação 2, onde $PI(x, y)$ é a função que descreve o grau de iteratividade do pesquisador “x” no projeto “y”. $Tar(x, y)$ é o total de tarefas do pesquisador “x” que foram realizadas no projeto “y”. $Contrib(x, y)$ é o total de contribuições que o pesquisador “x” realizou no projeto “y”. $Mat(x, y)$ representa o total de materiais que o pesquisador “x” disponibilizou para o projeto “y”. Por fim, $Public(x, y)$ representa o total de publicações geradas pelo pesquisador “x” no projeto “y”. Essas funções poderiam ser refinadas utilizando-se as competências de cada um destes itens, mas por questão de desempenho isto foi ignorado.

Equação 2 – Área do retângulo do pesquisador em um projeto.

$$PI(x, y) = \sum Tar(x, y) + \sum Contrib(x, y) + \sum Mat(x, y) + \sum Public(x, y)$$

2. Contexto Comunidade:

- a. Conjunto de informações utilizadas:

- Função do participante na comunidade – Caso seja um moderador, a importância do participante é aumentada.
 - Competências – Competências e grau destas competências que o participante possui e que são equivalentes às competências relevantes a comunidade. Aqui também é feito um comparativo do grau de cada competência pessoal do participante com o grau das competências requeridas pela comunidade. Assim, quanto maior for a proximidade do grau destas competências, maior será a importância da pessoa para a comunidade.
 - Contribuição – Mede a interatividade e apoio do participante para os demais membros da comunidade através da disponibilização de materiais, enquetes, eventos, “links” e notícias, além da participação em fóruns. Quanto maior for a contribuição da pessoa para a comunidade, maior é a sua importância.
- b. A cor do retângulo é determinada pela Equação 3, onde $CR(x, y)$ representa a relevância do pesquisador “x” na comunidade “y”. $CompPesq(x, y)$ representa o valor da competência do pesquisador “x”, que é similar a competência da comunidade “y”. $CompCom(y, x)$ representa o valor das competências da comunidade “y” que é similar a competência do pesquisador “x”. $Func(x, y)$ representa a importância do pesquisador “x” na comunidade “y”, $Func(x, y)$ será 50 caso o pesquisador (x) seja o moderador da comunidade, caso contrário será 0. Assim como justificado em projeto, o valor 50 foi escolhido por indicar uma variação perceptível na cor da escala das cores, também usamos duas funções distintas $CompPesq(x, y)$ e $CompCom(y, x)$ pelo fato de seus graus também variarem de “1 a 3” e “1 a 10”, respectivamente.

Equação 3 – Relevância (cor) do pesquisador em uma comunidade.

$$CR(x, y) = (\sum CompPesq(x, y) * CompCom(y, x)) + Func(x, y)$$

- c. O tamanho do retângulo é determinado pela Equação 4, onde $CI(x, y)$ representa o número de interações do pesquisador “x” na comunidade “y”. $Ev(x, y)$ representa o número de eventos que o pesquisador “x” participou na comunidade “y”. $Enq(x, y)$ representa o número de enquetes (criadas e respondidas) pelo pesquisador “x” na comunidade “y”. $Mat(x, y)$ representa o número materiais que o pesquisador “x” disponibilizou na comunidade “y”. $Link(x, y)$ representa o número de links sugeridos pelo pesquisador “x” na comunidade “y”. $Not(x, y)$ representa o número de notícias disponibilizadas pelo pesquisador “x” na comunidade “y”. Finalmente, $Forum(x, y)$ representa o número de participações em fóruns (tópicos criados e respondidos) do pesquisador “x” na comunidade “y”.

Equação 4 - Área do retângulo do pesquisador em uma comunidade.

$$CI(x, y) = \sum Ev(x, y) + \sum Enq(x, y) + \sum Mat(x, y) + \sum Link(x, y) + \sum Not(x, y) + \sum Forum(x, y)$$

3. Contexto Organizacional:

a. Conjunto de informações utilizadas:

- Função do participante: corresponde às funções que a pessoa executa nos projetos e comunidades em que participa. Caso seja um moderador ou gerente, sua importância na organização é incrementada.
- Competências – Competências e grau destas competências que a pessoa possui e que são relevantes a algum projeto ou comunidade existente no GCC. Ou seja, quanto maior for o número de comunidades e projetos em que a pessoa é relevante, maior será sua importância na organização.
- Contribuição – Corresponde a todas as contribuições que a pessoa efetuou nos projetos e comunidades de que participa.
- Rede de Relacionamento – O número de conexões que a pessoa possui com os demais integrantes do GCC. Partimos do

princípio de que uma pessoa muito “conectada”, com vários relacionamentos, pode ser de extrema importância em uma organização.

- b. A cor do retângulo é determinada pela Equação 5. Onde UR (x, y, z) representa a relevância do pesquisador “x” na organização, isto é, o somatório da importância de todos projetos “y” e comunidade “z” do pesquisador “x”. PR (x, y) representa a relevância do pesquisador “x” no projeto “y” e NumProj (x) o total de projetos que o pesquisador “x” participa. CR (x, z) representa a relevância do pesquisador “x” na comunidade “z” e NumComun (x) é o total de comunidades que o pesquisador “x” participa. Finalmente, RS (x, y, z) representa o total de pessoas que o pesquisador se relacionou em algum projeto ou comunidade. Caso o pesquisador se relacione com a mesma pessoa em diferentes projetos ou comunidades, isso só será contabilizado uma vez. PR e CR já foram calculados previamente nas Equações 1 e 3, respectivamente.

Equação 5 – Relevância (cor) do pesquisador na organização.

$$UR(x, y, z) = \frac{\sum PR(x, y)}{NumProj(x)} + \frac{\sum CR(x, z)}{NumComun(x)} + \frac{\sum RS(x, y, z)}{NumProj(x) + NumComun(x)}$$

- c. O tamanho do retângulo é determinado Equação 6, onde UI (x, y, z) representa o total de interações do pesquisador na organização, isto é, o somatório de interações em todos projetos e comunidade do pesquisador. Sendo PI (x, y) o número de interações do pesquisador “x” no projeto “y” e NumProj (x) o total de projetos que o pesquisador “x” participa. Por fim, CI (x, z) representa o número de interações do pesquisador “x” na comunidade “z” e NumComun (x) é o total de comunidades que o pesquisador “x” participa. PI e CI já foram calculados previamente, conforme visto nas Equações 2 e 4, respectivamente.

Equação 6 – Área do retângulo do pesquisador na organização.

$$UI(x, y, z) = \frac{\sum PI(x, y)}{NumProj(x)} + \frac{\sum CI(x, z)}{NumComun(x)}$$

Para facilitar o entendimento da abordagem, um exemplo é executado de forma gradativa no decorrer desta seção. Este exemplo fará uso das informações existentes nas tabelas de 6 a 12. As Tabelas 6, 7 e 8 ilustram as competências dos pesquisadores “P1”, “P3” e “P4” e seus respectivos graus. A Tabela 9 exhibe na primeira coluna as competências requeridas pelo projeto “Proj1”, na segunda coluna o grau de mínimo de conhecimento exigido em cada uma dessas competências e nas demais colunas o número de vezes que cada uma dessas competências foi utilizada em cada item do projeto. Por último, as Tabelas 10, 11 e 12 apresentam a participação dos pesquisadores P1, P3 e P4 no projeto Proj1, do qual P1 é gerente e irá sair. Na primeira coluna destas tabelas são listadas as competências que os pesquisadores e Proj1 apresentam em comum (CSCW – Trabalho Cooperativo Suportado por Computador, GC – Gestão do Conhecimento e BRI – Busca e Recuperação da Informação), enquanto as demais colunas exibem a utilização dessas competências pelo pesquisador em questão.

Tabela 6 – Competências do pesquisador P1 e seus respectivos graus.

	Baixo (valor 1)	Médio (valor 2)	Alto (valor 3)
CSCW			X
GC			X
BRI		X	
Bando de Dados			X

Tabela 7 – Competências do pesquisador P3 e seus respectivos graus.

	Baixo (valor 1)	Médio (valor 2)	Alto (valor 3)
CSCW			X
GC	X		
BRI		X	
Autômatos			X

Tabela 8 – Competências do pesquisador P4 e seus respectivos graus.

	Baixo (valor 1)	Médio (valor 2)	Alto (valor 3)
CSCW			X
GC			X
BRI		X	
Redes	X		

Tabela 9 – Competências do projeto Proj1 e sua utilização nos respectivos itens.

	Requeridas	Atividades	Contribuições	Materiais	Publicações
CSCW	7	12	9	4	9
GC	6	14	7	8	15
BRI	3	9	6	3	8
Comp Graf	5	6	5	2	0

Tabela 10 – Pesquisador P1 e sua participação no projeto Proj1.

	Atividades	Contribuições	Materiais	Publicações
CSCW	3	2	1	3
GC	7	3	3	7
BRI	3	4	2	3

Tabela 11 – Pesquisador P3 e sua participação no projeto Proj1.

	Atividades	Contribuições	Materiais	Publicações
CSCW	1	1	1	2
GC	2	1	3	0
BRI	3	1	2	3

Tabela 12 – Pesquisador P4 e sua participação no projeto Proj1.

	Atividades	Contribuições	Materiais	Publicações
CSCW	2	2	3	3
GC	3	3	2	1
BRI	2	4	2	3

A Figura 36 apresenta um exemplo de execução da etapa “Descobrir Contextos Críticos” na FBEE, ou seja, no nosso protótipo funcional da abordagem implementada no GCC. O usuário definiu o pesquisador P1 para ser substituído e selecionou o projeto Proj1, gerando o *Mapa de Árvores* apresentado. As cores e tamanhos dos retângulos do mapa em questão são calculados pelas fórmulas das Equação 1 e Equação 2, respectivamente. A fórmula da Equação 1 utiliza as informações contidas nas Tabelas 6, 9 e 10, além do fato de P1 ser gerente em Proj1, na determinação da cor do retângulo de P1. A primeira coluna da Tabela 10 exibe as competências que P1 e Proj1 possuem em comum, neste caso CSCW, GC e BRI, enquanto a Tabela 6 e a primeira coluna da Tabela 9 possuem os valores que são utilizados para o cálculo da cor. Logo $PR (“P1”, “Proj1”) = (7 \times 3 + 6 \times 3 + 3 \times 2) + 50 = 95$. Esse valor de PR é associado à tabela de cores no canto inferior esquerdo da Figura 36, o que atribuiu a cor verde clara ao retângulo referente ao pesquisador P1, indicando que ele é muito importante para o projeto. O tamanho do mesmo retângulo é calculado pela Equação 2 a partir das informações da Tabela 10, onde $PI (“P1”, “Proj1”) = 13 + 9 + 6 + 13 = 41$. Esse valor de PI gerou um retângulo cuja área é superior a dos demais retângulos, isso indica que a P1 é quem mais interage no projeto Proj1. Logo, com base na cor e tamanho do retângulo referente à P1, concluiu-se que a saída deste pesquisador causará grande impacto a Proj1.

Calculando os valores dos retângulos de P3 (Tabela 7, Tabela 9 e Tabela 11) e P4 (Tabela 8, Tabela 9 e Tabela 12), para confirmar a afirmação acima, temos: $PR (“P3”, “Proj1”) = 33$ e $PI (“P3”, “Proj1”) = 20$; $PR (“P4”, “Proj1”) = 45$ e $PI (“P4”, “Proj1”) = 30$. Logo, conforme mostrado na Figura 36, observamos que P1 apresenta uma coloração e área superior a P3 e P4, o mesmo vale para P4 em relação a P3.

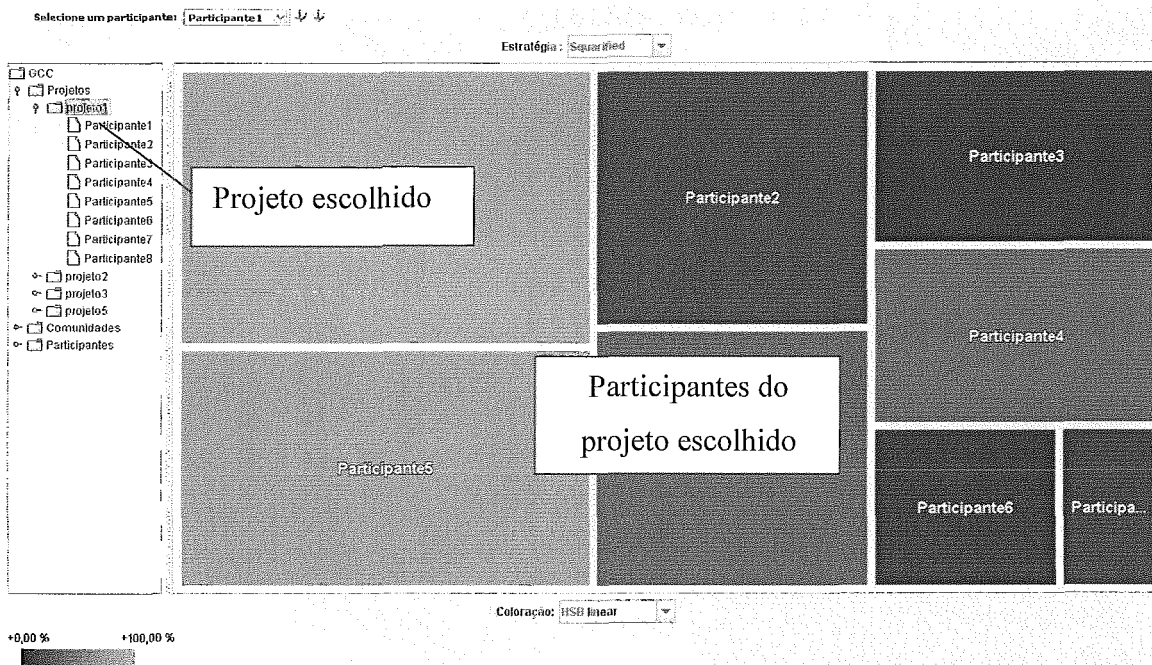


Figura 36 – Relevância das pessoas no projeto.

Conforme foi dito na seção 3.2.1.2, o *Mapa de Árvores* é uma técnica para visualizar informação hierárquica. O usuário visualiza de maneira rápida e interativa todos os contextos do pesquisador a ser substituído e determina quais projetos e comunidades serão mais prejudicados após a saída do mesmo, ou seja, define seu contexto de busca. Entretanto, a estrutura do *Mapa de Árvores* dificulta a exploração do aspecto semântico⁸ e dinâmico da informação. Com objetivo de suprir estas limitações, outras visualizações são utilizadas na BEE, as quais apresentaremos a seguir.

4.3.3. Localizar Substitutos Aptos

O objetivo desta etapa da BEE é indicar ao usuário possíveis pesquisadores substitutos a partir das competências do contexto selecionado. O uso dessas competências possibilita que a escolha do candidato substituto seja feita comparando-se o grau de conhecimento das pessoas em cada competência. Caso o usuário ainda esteja na dúvida

⁸ Nessa dissertação, consideramos o aspecto semântico da informação como sendo a comparação dos usuários a partir de suas competências.

sobre qual candidato escolher, outros critérios podem ser utilizados como fator de desempate, são eles: o perfil MBTI, o filtro de aprendizado numa competência e o filtro de interação com um tema de interesse.

Na primeira parte dessa etapa, o usuário tem a sua disposição um conjunto de competências relevantes a sua busca, devendo escolher a partir deste quais competências considera importante a um substituto. Lembramos que este conjunto corresponde às competências que são relevantes ao projeto, comunidade ou organização definido como o contexto de busca na etapa anterior. A partir daí é gerada uma visualização das competências escolhidas e seus respectivos graus. Essa visualização procura orientar a escolha de um substituto com base na discrepância do grau de conhecimento das pessoas em relação ao grau de conhecimento exigido pelo projeto ou comunidade. Ao compararmos pessoas a partir da diferença entre conhecimentos, estamos explorando o aspecto semântico da informação, categorizado no *espaço de informação de contexto* apresentado na seção 3.3. Logo, precisamos utilizar alguma técnica de visualização que facilite a percepção dessas diferenças.

Os gráficos de Teia ou Área foram as visualizações escolhidas para esta parte da busca. Esta técnica possibilita que a discrepância de conhecimentos seja percebida através da diferença entre as áreas das visualizações geradas, facilitando com isso o processo de tomada de decisões. O ideal é que uma pessoa *cubra* ao máximo a área que corresponde às competências dos projetos ou de comunidades, ou as competências que o usuário sugere como sendo essenciais para a continuação do projeto ou comunidade.

As Figura 37 e Figura 38 ilustram um exemplo de execução deste passo da abordagem através da FBEE, no qual as competências selecionadas do projeto Proj1 (BRI, GC e CSCW) e as pessoas que possuem todas essas competências são exibidas. A FBEE é dividida em duas partes, uma de consulta e outra de visualização. Na parte de consulta estão presentes:

1. O conjunto de competências de busca: este conjunto define as competências que podem ser utilizadas na busca, as quais correspondem as competências relevantes ao projeto ou comunidade selecionado, Figura 37 (a).
2. Somente as competências relevantes: essa opção esta relacionada às competências que são exibidas na visualização gerada. Se o usuário optar em

exibir as relevantes, somente as competências utilizadas na busca são exibidas. Caso opte por exibir todas, as competências não relevantes possuídas pelos candidatos substitutos também são exibidas, Figura 37 (b).

3. O tipo de visualização a ser utilizada: o usuário pode optar por comparar o grau das competências através da visualização de teia ou área, Figura 37 (c).
4. As visualizações auxiliares: essas visualizações são utilizadas quando o usuário considerar que somente a discrepância de conhecimento dos candidatos não é suficiente para decidir qual pessoa escolher, Figura 37 (d). Esses itens correspondem a Similaridade MBTI; Interesse do Usuário e Utilização de Competências e são detalhados nas seções 4.3.3.1, 4.3.3.2, 4.3.3.3.
5. O peso dos itens utilizados: esse peso corresponde à importância que um determinado item tem na valoração do grau da competência, Figura 37 (e). O cálculo desse grau é explicado a seguir.

A parte de visualização da FBEE, também é composta de duas partes, na parte superior, Figura 37 (i), são visualizados as competências do contexto definidas para busca, as da pessoa a ser substituída e de cada um dos candidatos recomendados separadamente, os quais podem ser vistos nas Figuras 37 (f), (g) e (h), respectivamente. A parte inferior, Figura 37 (j), faz o mesmo tipo de comparação, porém todos os candidatos são visualizados na mesma tela ao mesmo tempo. Essa divisão foi realizada pelo fato de que quando muitas competências são exibidas ao mesmo tempo, a visualização inferior torna-se densa Figura 37 (l), o que pode dificultar a percepção das discrepâncias.

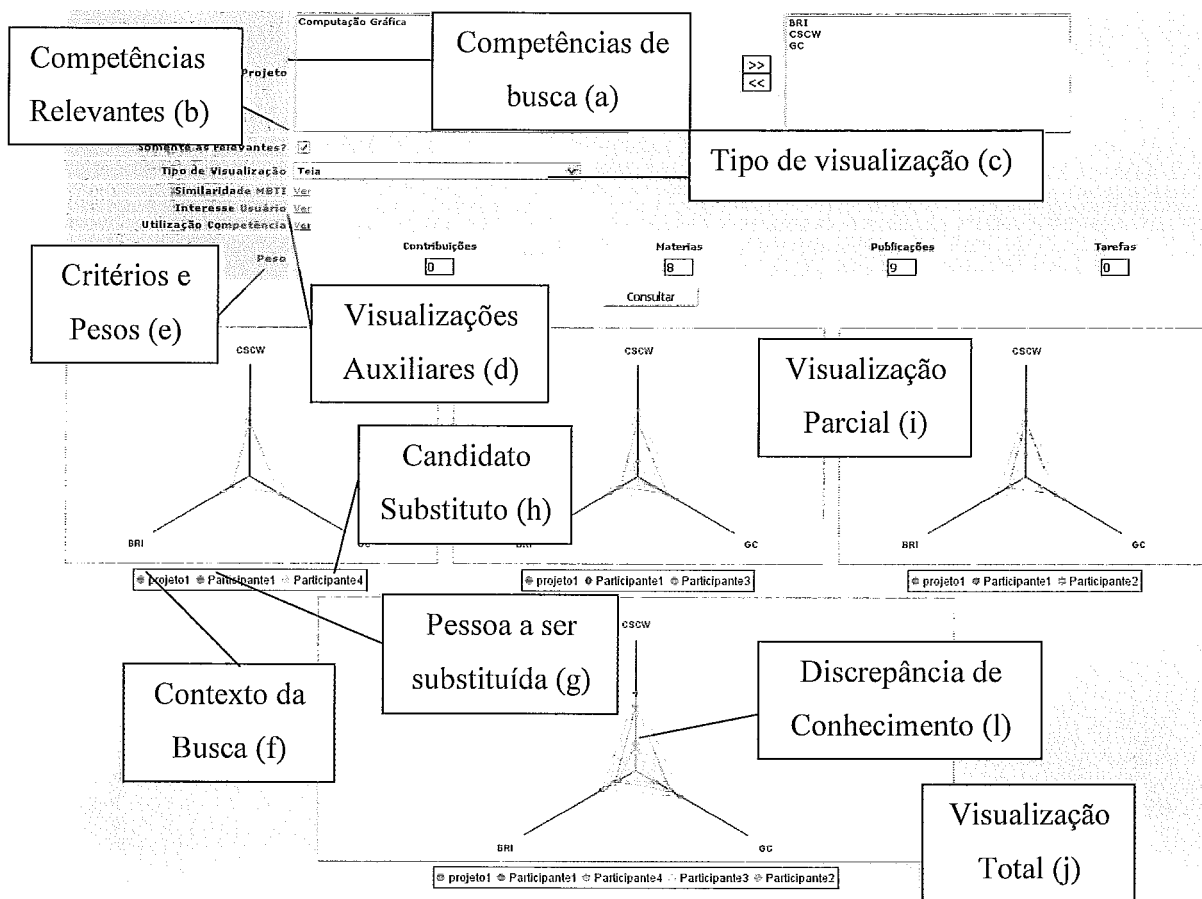


Figura 37 – Visualização de Teia gerada na substituição de uma pessoa no contexto de projeto.

As mesmas considerações feitas para visualização de Teia (Figura 37) valem para visualização de Área (Figura 38), a única diferença é a maneira que a informação é apresentada ao usuário e sua aplicabilidade. A visualização de Área é recomendada para os casos em que estamos avaliando um pequeno número de competências, pois com muitas competências existe uma grande sobreposição de gráficos, o que dificulta a análise. Nesse caso, a visualização de Teia é recomendada, pelo fato do número de eixos ser equivalente ao número de competências, a sobreposição de áreas fica mais perceptível, facilitando a identificação da diferença de áreas entre as mesmas.

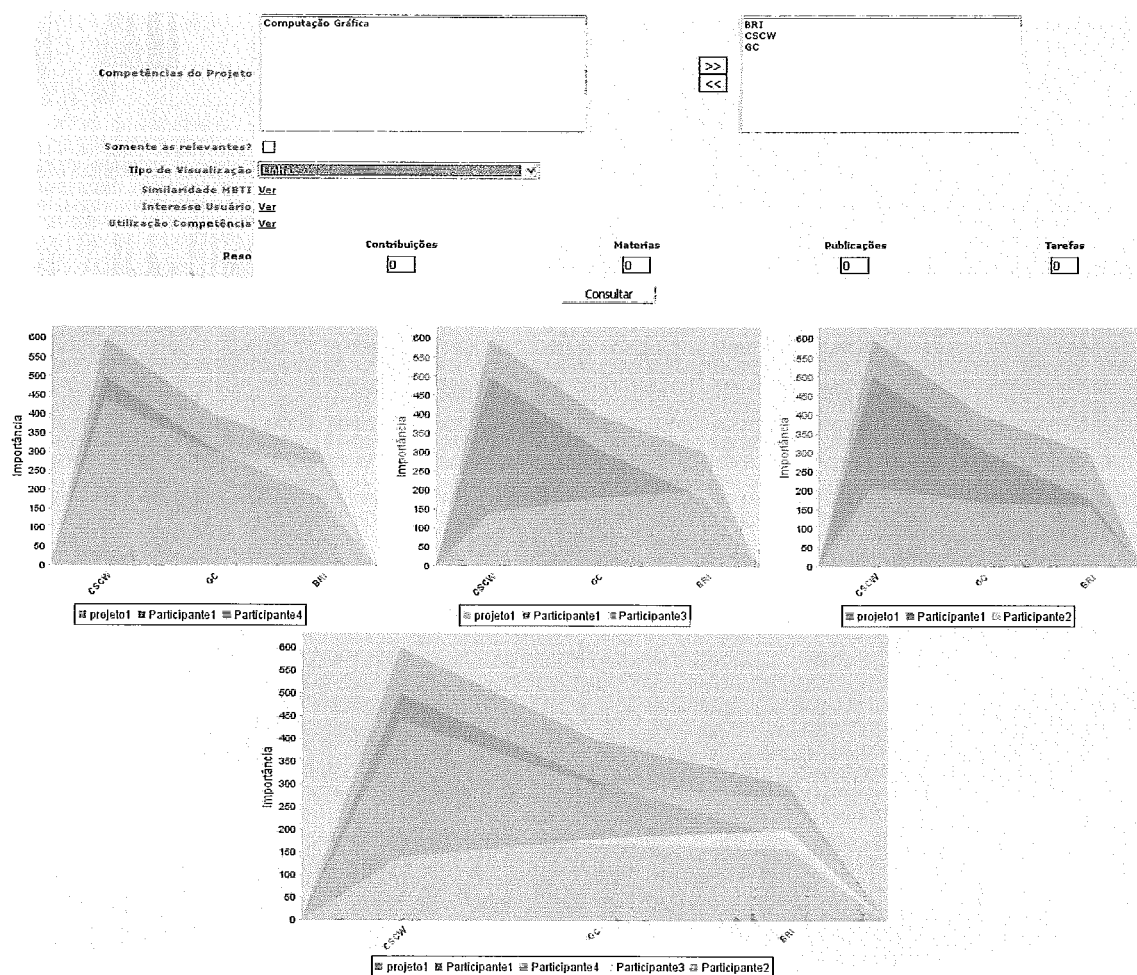


Figura 38 – Visualização de Área gerada na substituição de uma pessoa no contexto de projeto.

Outro fator importante é como calcular o grau das competências. Na BEE, o grau das competências de projetos e comunidades é calculado a partir do grau da competência no contexto de busca, valor entre 1 e 10, juntamente com os itens relacionados a essa competência no mesmo contexto. Já o grau das competências dos candidatos substitutos e da pessoa a ser substituída é definido a partir do nível de conhecimento (alto, médio e baixo, respectivamente 3, 2 e 1) da pessoa na competência, mais os itens relacionados a esta competência e que foram disponibilizados pela mesma no contexto em questão.

A valoração do grau das competências não está integrado ao módulo de Busca de Competências do GCC, o qual foi detalhado na seção 4.2.1.2. O principal motivo é que a mensuração das competências realizada no módulo de Busca de Competências faz uma

agregação utilizando várias critérios, os quais foram mostrados na Figura 31, sem considerar contextos específicos. A necessidade de se utilizar outra maneira de se valorar o grau das competências na BEE é para atender esta abordagem, que diferencia a busca através dos diferentes contextos.

Um detalhe importante na determinação do grau das competências consideradas refere-se a livre atribuição de pesos para cada tipo de item utilizado na valoração. Esta abertura no critério de definição de pesos é importante porque o usuário que busca pessoas com determinada competência, ou conjunto de competências, pode priorizar a experiência profissional (via Atividades executadas no projeto) ou participação acadêmica (via Publicação), por exemplo, bastando dar maior peso para os itens escolhidos. Os itens e fórmulas utilizados em cada contexto de busca (projeto, comunidade e organização) para determinar o grau de cada competência são detalhados a seguir:

1. Contexto Projeto:

a. Conjunto de informações utilizadas:

- Competências requeridas pelo projeto – o uso dessas competências limita o espaço de busca do usuário ao conjunto de competências que o pesquisador a ser substituído possui e que são realmente relevantes ao projeto. Essa redução propicia ao usuário recomendações mais concisas, aumentando assim a chance de se encontrar um bom substituto. Lembramos que o valor dessas competências variam entre 1 e 10.
- Competências dos pesquisadores candidatos – às vezes somente as competências relevantes ao projeto não são suficientes para encontrar um substituto, sendo assim, a BEE possibilita que as competências não relevantes ao projeto também sejam analisadas, caso seja desejo do usuário. Lembramos que o valor das competências de um pesquisador recebem grau alto, médio e baixo cujos valores são 3, 2 e 1, respectivamente.
- Itens do contexto – são os itens correspondentes às interações dos participantes do projeto, tais como: contribuições, materiais, tarefas e publicações, já explicados na seção anterior. Esses itens

são utilizados na valoração do grau de uma competência. Partimos do pressuposto que quanto maior o número de itens que um participante possui relacionado a uma competência, maior é seu conhecimento na mesma.

Todos esses atributos são utilizados nas fórmulas das Equações 7 e 8.

- b. O grau das competências do projeto é determinado pela Equação 7, onde $PC(x, y)$ é o grau da competência “x” no projeto “y”, representado a partir do valor associado e itens relacionados à mesma. $CompProj(x, y)$ representa o valor diretamente associado a competência “x” (1-10) no projeto “y”, essa atribuição é realizada durante o cadastro do projeto. $Item(x, y, z)$ representa o total de itens do tipo “z” relacionados à competência “x” no projeto “y”. Conforme mostrado anteriormente, nós consideramos 4 tipo de itens (contribuições, materiais, tarefas e publicações) no contexto projeto, sendo assim Z irá variar de 1 a 4. Por fim, $P(z)$ é o peso (1 - 10) que o item “z” terá no calculo do grau da competência “x” no projeto “y”. Essa variação de peso permite ao usuário valorar as competências de acordo com os itens que considera mais importante.

Equação 7 – Grau das competências do projeto.

$$PC(x, y) = CompProj(x, y) + \sum_{Z=1}^4 (Item(x, y, z) \times P(z))$$

- c. O grau das competências dos candidatos pertencentes ao projeto é determinado pela Equação 8, em que $PPC(w, x, y)$ é o grau da competência “x” do candidato “w” no projeto “y”, representado a partir do valor e itens associados à mesma. $CompPesqProj(w, x)$ representa o valor diretamente atribuído a competência “x” (1, 2 ou 3 respectivamente baixo, médio ou alto) pelo pesquisador “w” quando a cadastrou em seu perfil. Como essa valoração é muito subjetiva, por exemplo, duas pessoas podem ter diferentes percepções do que

representa a atribuição média para uma competência, os itens são utilizados para melhorá-la. O Item (w, x, y, z) representa o total de elementos do tipo “z” relacionados à competência “x” do projeto “y” e que estão associados ao pesquisador “w”. Conforme explicados no item b, 4 tipos de itens são considerados, portanto Z varia de 1 a 4. Finalmente, P(z) é o peso que o item “z” tem no cálculo do grau da competência.

Equação 8 – Grau das competências dos pesquisadores candidatos pertencentes ao projeto.

$$PPC(w, x, y) = \text{CompPesqProj}(w, x) + \sum_{z=1}^4 (\text{Item}(w, x, y, z) \times P(z))$$

2. Contexto Comunidade:

a. Conjunto de informações utilizadas:

- Competências requeridas pela comunidade – são as competências relevantes às comunidades, ou seja, os conhecimentos debatidos na mesma. Essas competências recebem um valor de 1 a 10, que é atribuído na criação da comunidade.
- Competências dos pesquisadores candidatos – são as competências que os participantes da comunidade possuem, os quais variam de 1 a 3 (baixo, médio e alto).
- Itens do contexto – são os itens correspondentes às interações dos participantes na comunidade, tais como: enquetes, eventos, participações em fóruns, “links”, materiais e notícias, também explicados na seção anterior. Esses itens são utilizados na valoração do grau de uma competência.

Todos esses atributos são utilizados nas fórmulas das Equações 9 e 10.

- b. O grau das competências da comunidade é determinada pela Equação 9, onde CC (x, y) é o grau da competência “x” na

comunidade “y”, assim como em projetos, o grau de cada competência da comunidade é definido a partir de seu valor associado juntamente com os itens relacionados a mesma. CompCom (x, y) representa o valor diretamente (1-10) associado a competência “x” da comunidade “y” durante sua criação. Item (x, y, z) representa o total de itens do tipo “z” relacionados à competência “x” da comunidade “y”. No contexto de comunidade são considerados 6 tipos de itens (enquetes, eventos, participações em fóruns, “links”, materiais e notícias), sendo assim Z varia de 1 a 6. Por fim, P(z) é o peso que um item “z” tem no cálculo do grau da competência “x”.

Equação 9 – Grau das competências da comunidade.

$$CC(x, y) = \text{CompCom}(x, y) + \sum_{z=1}^6 (\text{Item}(x, y, z) \times P(z))$$

- c. O grau das competências dos candidatos pertencentes a comunidade é determinado pela Equação 10, em que PCC (w, x, y) é o grau da competência “x” do candidato “w” na comunidade “y”, representado a partir do valor e itens associados a esta competência. CompPesqCom(w, x) representa o valor diretamente associado a competência “x” (1, 2 ou 3 respectivamente baixo, médio ou alto) pelo pesquisador “w” quando cadastrou essa competência em seu perfil. Item (w, x, y, z) representa o total de itens do tipo “z” relacionados à competência “x” da comunidade “y” e que estão associados ao pesquisador “w”. Esses itens são utilizados para melhorar a valoração do grau das competências dos pesquisadores pertencentes à comunidade, visto que sua atribuição é muito subjetiva. Como existem 6 tipos de itens associados ao pesquisador, o valor de Z vai variar de 1 a 6. Finalmente, P(z) é o peso que o item do tipo “z” tem no cálculo do grau da competência.

Equação 10 – Grau das competências dos pesquisadores candidatos pertencentes à comunidade.

$$PCC(w, x, y) = \text{CompPesqCom}(w, x) + \sum_{z=1}^6 (\text{Item}(w, x, y, z) \times P(z))$$

3. Contexto Organizacional:

a. Conjunto de informações utilizadas:

- Competências do pesquisador a ser substituído – nesse contexto o conjunto de competências da pessoa a ser substituída limita o espaço de busca, diferentemente dos contextos anteriores em que esse espaço é limitado pelas competências do contexto selecionado. Essa mudança ocorre porque quando o usuário opta por buscar um substituto dentro da organização, o fato dele pertencer ao mesmo projeto ou comunidade da pessoa a ser substituída torna-se irrelevante. Logo, a similaridade de conhecimentos entre o par substituto e substituído passa a predominar. Como estamos considerando competências pessoais, a atribuição passa a ser baixo, médio ou alto (1, 2 ou 3).
- Competências dos pesquisadores candidatos – são as competências dos candidatos a substitutos, as quais também recebem valores alto, médio ou baixo.
- Itens do contexto organizacional: são todos os itens que as pessoas possuem e que estão relacionados à competência em questão. Conforme mostrado anteriormente, esses itens são disponibilizados em dois tipos de contexto: projetos (contribuições, materiais, tarefas e publicações) e comunidades (enquetes, eventos, participações em fóruns, links, materiais e notícias). Logo, a partir do momento em que buscamos um substituto numa organização, qualquer item que ele possua relacionado à competência deve ser considerado.

Todas esses atributos são utilizados na Equação 11.

- b. O grau das competências dos candidatos é determinado pela Equação 11, onde $UC(w, v, x, y)$ é o grau da competência “x” pertencente ao pesquisador “w” em um projeto de “y” ou em uma comunidade de “v”. Nesse caso, “y” representa o conjunto de projetos na organização que o pesquisador atua, por exemplo, se o pesquisador atua em 3 projetos $y = 3$. Já “v” corresponde ao conjunto de comunidades que o pesquisador participa. $PCC(w, x, y)$ é o grau da competência “x” pertencente ao pesquisador “w” projeto “y”. Enquanto $PPC(w, x, v)$ é o grau da competência “x” pertencente ao pesquisador “w” na comunidade “v”. PPC e PCC já foram explicados anteriormente.

Equação 11 – Grau das competências dos pesquisadores candidatos pertencentes à organização.

$$UC(w, v, x, y) = \sum_0^y PCC(w, x, y) + \sum_0^v PPC(w, x, v)$$

Continuando o exemplo da seção anterior, ou seja, a substituição de P1 em Proj1. Os graus das competências do projeto e de seus respectivos participantes são determinados pelas Equações 7 e 8, respectivamente. Por exemplo, considerando o cálculo da competência GC para Proj1 e P1, sendo todos os itens com peso 1 temos: $PC(“GC”, “Proj1”) = 6 + (14 \times 1 + 7 \times 1 + 8 \times 1 + 15 \times 1) = 50$ e $PPC(“P1”, “GC”, “Proj1”) = 3 + (7 \times 1 + 3 \times 1 + 3 \times 1 + 7 \times 1) = 23$. Os valores de todas as competências do respectivo projeto e seus integrantes são apresentados na Tabela 13. A primeira coluna contém o projeto e participantes considerados. A segunda coluna contém as tabelas utilizadas como fonte de informação no cálculo das competências. A terceira coluna apresenta as equações utilizadas na determinação do grau. Por fim, as duas últimas colunas contêm a competência calculada e seu respectivo valor.

Tabela 13 – Cálculo do grau das competências do projeto e seus participantes.

Projeto/Participantes	Tabelas	Equação	Competências	Valor
Proj1	9	7	CSCW	41
			GC	50
			BRI	29
P1	6 e 10	8	CSCW	12
			GC	23
			BRI	14
P3	7 e 11	8	CSCW	8
			GC	7
			BRI	11
P4	8 e 12	8	CSCW	13
			GC	12
			BRI	13

Com base nesses dados, observa-se que a área do gráfico de Proj1 > P1 > P4 > P3, para confirmar essa afirmação basta retornar as Figura 37 e Figura 38. Logo, com base nessas visualizações a FBEE recomenda P4 como o substituto apropriado para P1. Caso P4 esteja indisponível, P3 também seria uma boa escolha. A seguir, algumas considerações sobre as visualizações de Teia e Área são efetuadas:

1. Caso a FBEE encontre mais de cinco pesquisadores no contexto considerado, somente os cinco mais similares são exibidos na visualização. Caso contrário, a parte inferior da visualização ficaria muito confusa.
2. As visualizações (área ou teia) que são geradas na parte superior estão ordenadas por ordem crescente de similaridade da esquerda para a direita.
3. A visualização de teia deve ser utilizada somente quando houver mais de duas competências e menos de oito. Com uma ou duas competências o gráfico gerado será uma linha, o que não acrescenta nada na percepção da diferença semântica entre os pesquisadores. Acima de oito competências, a visualização de teia torna-se confusa.

4. Ambas as visualizações ilustram a diferença semântica entre os pesquisadores, o que muda é a forma de se observar essa diferença.
5. O primeiro elemento da legenda refere-se ao contexto analisado e o segundo refere-se sempre ao pesquisador a ser substituído. Finalmente, os elementos restantes referem-se aos possíveis substitutos. No caso do contexto organizacional, a legenda será composta de dois elementos, o substituído e candidatos a substituto.
6. Os limites do gráfico são iguais à competência de maior grau, por exemplo, se a competência de maior grau tiver valor 300, o limite do gráfico será 300.

Entretanto, caso ocorra uma situação de indicações mútuas nessa etapa, outros critérios podem auxiliar a escolha do pesquisador substituto, são eles: a similaridade MBTI, o filtro de aprendizado numa competência e o filtro de interação com um tema de interesse, os quais são apresentados nas seções 4.3.3.1, 4.3.3.2, 4.3.3.3., respectivamente.

4.3.3.1. Similaridade MBTI

O MBTI (MYERS, 1980), Inventário Tipológico de Myers-Briggs, foi desenvolvido por Katherine Briggs e por sua filha Isabel Myers nos EUA. Baseado na psicologia de Carl Gustav Jung (JUNG, 1971), psiquiatra suíço que criou a teoria dos tipos psicológicos, o MBTI é composto por um questionário cujos resultados apontam as diferenças existentes entre as pessoas. O MBTI proporciona uma medida da personalidade considerando oito tendências que as pessoas usam em diferentes momentos, o conhecimento dessa diversidade pode auxiliar as pessoas a se entenderem melhor, essa é a principal função do MBTI. Essas características permitem a rápida identificação do perfil das pessoas e contribuem para a formação de equipes, razão pelas quais esse perfil foi adicionado ao GCC.

O MTBI estabelece 16 tipos para demonstrar as tendências de cada indivíduo. Esses tipos são formados mediante as diferentes combinações possíveis entre quatro pares de dimensões (preferências) existentes. As pessoas podem preferir concentrar sua atenção na Extroversão (E) ou Introversão (I); obter informações pela Sensação(S) ou Intuição(N); tomar decisões pelo Pensamento (T) ou Sentimento (F); e se orientar em relação ao mundo

exterior pelo Julgamento (J) ou Percepção (P). O tipo do perfil de uma pessoa é determinado através de respostas a um questionário.

Segundo KROEGER & THUESEN (1988) e KEIRSEY & BATES (1984), não existe entre os 16 tipos psicológicos um tipo que seja melhor do que outro, porém existem tipos mais adequados para uma função ou atividade do que outros. A Tabela 14 apresenta 16 tipos de perfil possíveis.

Tabela 14 – Tipos de perfis de acordo com a tipologia MBTI MYERS (1995) e KEIRSEY & BATES (1984).

ISTJ	ISFJ	INFJ	INTJ
ISTP	ISFP	INFP	INTP
ESTP	ESFP	ENFP	ENTP
ESTJ	ESFJ	ENFJ	ENTJ

O uso da similaridade MBTI é diferente dos demais critérios, pois neste caso estamos considerando as características do perfil psicológico dos pesquisadores e não suas competências. Visto que, muitas vezes a identificação da personalidade é igualmente importante quanto saber o nível de conhecimento de uma pessoa, para a realização de uma parceria ou projeto de pesquisa. Às vezes torna-se mais conveniente ter como membro de equipe uma pessoa menos especializada em um assunto, mas interativa e comunicativa, do que um gênio introvertido.

O gráfico de força (Figura 39) foi a visualização escolhida para representar a similaridade do perfil MBTI. Nessa visualização, a pessoa a ser substituída, Figura 39 (a), fica ao centro e os candidatos substitutos Figura 39 (b), ficam a sua volta, quanto mais similar o perfil dos candidatos substituto em relação a pessoa a ser substituída, maior a proximidade entre os mesmos. A idéia para criar esse gráfico foi decorrente da visualização utilizada no aplicativo SINNZEUG, conforme ilustrado na Figura 17, em que “sites” com semântica similar a de uma palavra inserida na tela aproximam-se da mesma. A visualização MBTI pode ser utilizada em dois momentos:

1. Quando as visualizações para análise de similaridades de competências indicaram mais de um substituto, a visualização MBTI pode ser utilizada como critério de desempate.

- Quando as características do perfil são mais importantes que as competências, essa visualização por si só pode indicar o pesquisador a ser escolhido.

A Figura 39 apresenta um exemplo, através da FBEE, de utilização do *gráfico de forças* na visualização de similaridades de perfil MBTI. Nesta visualização podemos escolher como referência o perfil que mais se assemelha com o da pessoa a ser substituída. Neste caso, quanto mais similar, mais próximo a pessoa fica do membro a ser substituído. Uma outra possibilidade é a escolha de certas características específicas. Nessa opção um membro hipotético com características escolhidas é gerado e a proximidade em relação ao mesmo é calculada. Essa alternativa é válida quando estamos buscando pessoas que apresentem características de perfil destoantes das da pessoa a ser substituída.

Uma descrição das características similares que o candidato substituto possui é exibida ao selecioná-lo, por exemplo, na Figura 39 (b) o participante 8 possui somente duas características similares a P1. O participante 5, no caso o mais próximo, certamente apresenta todas as quatro características de seu perfil similar a P1. Logo, se estamos buscando uma pessoa que apresente as mesmas características de P1, então P5 é o indicado.

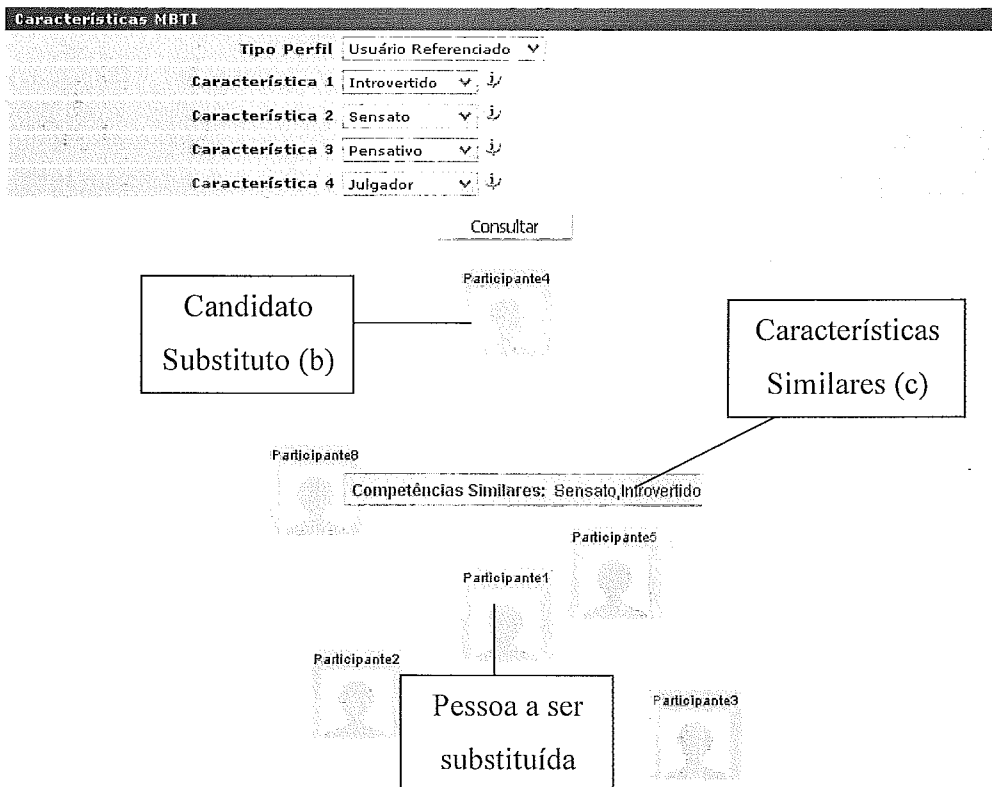


Figura 39 – Visualização para Similaridade MBTI.

Ao apresentarmos as similaridades de perfil MBTI entre candidatos substitutos e a pessoa a ser substituída através de uma visualização que realce essa proximidade, estamos evitando que o usuário tenha que analisar relatórios referentes ao perfil de cada um dos candidatos, reduzindo assim o esforço cognitivo requerido ao mesmo. A seguir é apresentado outro critério de desempate para o caso de recomendações similares.

4.3.3.2. Filtro de Aprendizado e Filtro de Interação com um Tema de Interesse

Pode acontecer que durante o cálculo da similaridade de competências, os pesquisadores recomendados apresentem o mesmo nível de conhecimento nas mesmas. Uma maneira de optar pela pessoa correta é descobrir quais pesquisadores trabalharam mais recentemente com uma competência, ou o quão crescente este trabalho é. Portanto, nessa dissertação, filtro de aprendizado corresponde ao quanto uma pessoa aprendeu sobre um determinado assunto (competência) num período de tempo.

Nessa abordagem, o nível de conhecimento em um determinado assunto é aferido através da quantidade de itens que uma pessoa possui, em seu contexto de atuação, no assunto em questão. No contexto do projeto são considerados publicações, contribuições, materiais criados ou sugeridos (apenas no caso de materiais) por um pesquisador e que estão relacionados a uma competência. São também levadas em considerações as atividades executadas e concluídas com sucesso pelo pesquisador que possuem como pré-requisito esta competência. No contexto de comunidade são considerados todos os tipos de interação – como notícias, enquetes, eventos, “links”, materiais e fóruns – sobre um determinado assunto (competência). Por último o contexto da organização considera todos os itens dos dois contextos anteriores. Logo, precisamos apresentar esses itens de modo sua utilização em um período de tempo seja facilmente perceptível.

O *gráfico de linha* foi a visualização escolhida para essa etapa da abordagem. Esse tipo de gráfico é representado em duas dimensões, com base na representação cartesiana dos pontos no plano. Sendo a melhor opção para mostrar mudanças num grupo de valores em função do tempo (para longos períodos de tempo), especialmente quando se deseja enfatizar e analisar a mudança de taxas ao longo do tempo. Além de possibilitar a

visualização de grandes grupos de dados simultaneamente. Como estamos interessados em análises quantitativas, ou seja, o número de contribuições de um pesquisador relacionadas a um assunto, essa visualização tornou-se suficiente.

A Figura 40 mostra um exemplo da visualização de *filtro de aprendizado em um assunto* executado através da FBEE. O nome da competência a ser analisada está presente no título do gráfico gerado. O eixo das abscissas (X) contém os itens do contexto que são utilizados na mensuração da competência. O eixo das ordenadas (Y) contém a quantidade de cada um desses itens. Cada linha de cor diferente representa um pesquisador descrito na legenda. Por fim, o grau final de conhecimento numa dada competência é indicado na coluna *total* do eixo Y. Existe também a possibilidade de limitar o período de busca, como por exemplo, priorizar as pessoas com maior quantidade de contribuição no último ano. Neste exemplo, o *Participante1*, representado pela linha vermelha, possui maior nível de envolvimento com a competência *gestão do conhecimento*.

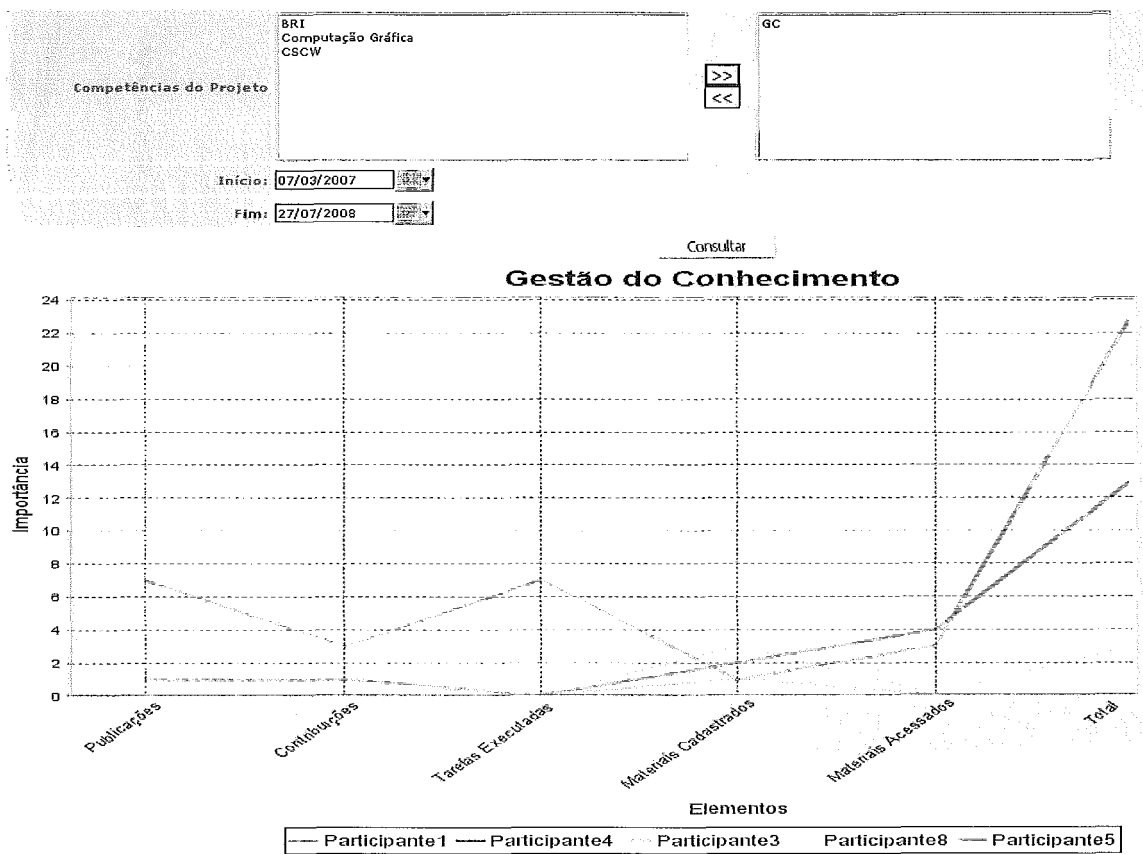


Figura 40 – Gráfico utilizado na visualização do nível de conhecimento de uma competência no contexto projeto.

A análise temporal provida por essa visualização é sua principal vantagem. A análise, a partir de intervalos de tempos, possibilita ao avaliador identificar o recente grau de envolvimento dos candidatos substitutos com o assunto. Isso facilita o processo de tomada de decisão, pois quanto maior for o envolvimento recente com um assunto, maiores são as chances do candidato estar atualizado e de aceitar o convite para substituir a pessoa que está abandonando a organização.

A análise do filtro de interação com um tema de interesse é equivalente à análise do filtro de aprendizado num assunto, ou seja, usa os mesmos itens e visualização. A diferença é apenas o conceito que adotamos para competência e interesse. Competência é um assunto que o pesquisador domina, algo que ele já pesquisou ou trabalhou, e no caso possui alguma “expertise”. Interesse representa algum conhecimento que ele *não* detém, mas quer aprender, pesquisar ou trabalhar. À medida que um pesquisador progride seu aprendizado em um determinado interesse, este pode virar uma competência. Portanto, o filtro de interação com um tema de interesse mostra o quanto uma pessoa interagiu com um tema do seu interesse.

Considerando que na busca por um substituto, procura-se uma pessoa que apresenta menor discrepância em uma competência, o uso do critério interesse seria incoerente. Porém, a análise de interesses tem grande utilidade quando estamos procurando novos estagiários – entende-se por estagiário uma pessoa interessada em aprender um novo assunto – ou novos membros para uma comunidade virtual. PREECE (2000) define comunidades virtuais como um local onde pessoas com interesses comuns se encontram, entram em contato e trocam idéias e informações umas com as outras através de uma rede “online”, como a Internet. Um exemplo do uso dessa abordagem na recomendação de comunidades virtuais pode ser visto em KREJCI *et al.* (2008a), na qual as competências de pesquisadores brasileiros são mineradas através da ferramenta S-Miner e posteriormente pesquisadores com pelo menos uma competência similar são apresentados pela técnica *Mapa de Árvores*.

Outro diferencial dessa visualização, em relação as anteriores (Árvore de Mapas, Gráfico de Força e Gráfico de Teia ou Área), é que pela primeira vez as informações avaliadas são apresentadas separadamente e têm seu fluxo analisado quantitativamente, cobrindo assim o terceiro e último espaço de informação (espaço de informação dinâmico).

Logo, as visualizações de filtro (Gráfico de Linha) juntamente com as de similaridade (Gráfico de Força e Gráfico de Teia ou Área) complementam a visualização utilizada na exibição da informação hierárquica (Mapa de Árvores), possibilitando assim a obtenção de resultados mais precisos.

Quanto aos aspectos técnicos, a implementação do protótipo pode ser dividida em duas partes:

1. A primeira parte refere-se à parte lógica da abordagem, ou seja, ao armazenamento e tratamento dos dados. Essa parte foi construída utilizando a arquitetura cliente-servidor, o servidor “web” IIS (Internet Information Services), ASP como linguagens de programação e o banco de dados MS-SQL Server 2000. Vale lembrar que esta é a plataforma de desenvolvimento do GCC, e como este protótipo foi implementado nesse ambiente, obedeceu aos mesmos critérios de implementação. A base de dados utilizada pela FBEE pode ser vista no Anexo II, respectivamente nas Figura 41, Figura 42 e Figura 43.
2. A segunda parte diz respeito às visualizações aplicadas na exibição dos dados. Como a linguagem ASP não provê suporte a bibliotecas gráficas, essas visualizações foram implementadas na linguagem de programação orientada a objetos Java (GOSLING *et al.*, 2005) e apresentadas na através de “Applets”. Além das ótimas características – tais como: portabilidade, escalabilidade e ser multiplataforma – essa linguagem apresenta uma grande quantidade de bibliotecas programadas disponível na “web” o que facilita o reuso de códigos já implementados, minimizando assim o tempo de programação. As visualizações de Mapa de Árvores e Gráfico de Força foram implementadas através da biblioteca gráfica *Prefuse* (PREFUSE, 2008), a qual está sendo muito utilizada na produção de visualizações interativas devido sua alta performance e simplicidade de manipulação (HEER *et al.*, 2005). As outras visualizações – Gráfico de Área, Linha e Teia – foram implementadas com a biblioteca *JFreeChart*, por ser uma biblioteca específica para a construção de gráficos (JFREECHART, 2008), a quantidade exemplos encontrados, somados a sua excelente documentação possibilitou a rápida implementação das visualizações citadas.

A seguir é apresentado um estudo comparativo da FBEE frente aos SREs apresentados na seção 3.5.

4.4. Comparação com Trabalhos Correlatos

Na literatura existe um grande número de sistemas de computadores (sistemas de busca de especialistas) que procuram facilitar a percepção do que uma pessoa sabe, isto é: suas competências, formação, áreas de interesse etc. A partir do momento que todo capital intelectual de uma pessoa torna-se perceptível a outras, maiores são as chances de existir colaboração, indicações para solução de problemas e indicação de substitutos. Apesar desses sistemas possuírem o mesmo objetivo, tornar o capital intelectual de uma pessoa perceptível, eles possuem características diferentes. Objetivando facilitar a classificação desses sistemas, YIMAM-SEID & KOBASA (2003) propuseram o *Modelo de Domínio Intuitivo para Sistemas de Busca de Especialista*. Este modelo é utilizado para comparar a FBEE com as demais ferramentas citadas anteriormente na seção 3.5.

Segundo BIRK (1997), um *Modelo de Domínio Intuitivo* é uma classificação de facetas que sumarizam as conclusões do domínio analisado e o descreve através de termos comuns a profissionais de “software”. Cada *faceta* no modelo de domínio é representada por um *fator de domínio* (atributos compartilhados pelos sistemas). Cada fator de domínio é definido por um conjunto de valores discretos (*valores possíveis*). Uma faceta de um domínio concreto é caracterizada através de um ou mais valores (*valores atuais*) provenientes do conjunto de valores possíveis. Um fator de domínio, juntamente ao seu valor atual, é chamado *característica do domínio*. YIMAM-SEID & KOBASA (2003) identificaram sete facetas para o domínio correspondente aos *sistemas de recomendação de especialistas*, conforme mostrado na Tabela 15, onde cada coluna representa uma faceta e seus possíveis valores, os quais representam possíveis implementações da respectiva faceta. Logo, os *sistemas de busca de especialistas* podem ser classificados através dos valores de cada faceta. A seguir explicamos cada uma das sete facetas:

1. Forma de reconhecer o especialista: o reconhecimento de um especialista pode ser realizado de forma explícita ou implícita. O reconhecimento explícito é feito com base nas competências declaradas pelo próprio especialista ou em sua posição profissional. Na implícita, as competências do especialista são

determinadas dinamicamente pelo sistema. Estas competências podem ser mineradas a partir de documentos dos próprios especialistas ou dos quais eles são co-autores ou são citados, de projetos que eles participam, dos serviços utilizados freqüentemente por eles e de suas páginas pessoais.

2. Operações de extração de indicadores de competências: as técnicas que são utilizadas na extração de competências podem ser dependentes ou não do domínio de conhecimento. Nas técnicas dependentes do domínio, determina-se previamente o local onde os indicadores de competências serão buscados, como por exemplo, o sistema desenvolvido por VIVACQUE & LIEBERMAN (2000) em que as habilidades em Java de um programador são extraídas a partir dos códigos por ele produzido. Nas não-dependentes de domínio, a busca por indicadores pode ser feita através de palavras-chave, freqüência de uso de uma determinada funcionalidade de um “software”, proximidade entre documentos etc.
3. Construção do perfil: o perfil do especialista pode ser gerado dinamicamente durante a consulta, ou construído e armazenado através do uso de agentes pessoais ou modelos agregados. Sistemas que constroem os perfis dos especialistas a serem buscados durante a consulta, utilizam os termos da consulta como base de construção do perfil. Nos sistemas baseados em agentes, os agentes pessoais possuem duas tarefas, a construção do perfil do especialista e a localização de outros especialistas. Por último, na abordagem que associa os especialistas a modelos agregados, os especialistas são associados a diferentes representações de competências que podem ser ontologias e estruturas organizacionais.
4. Mecanismos de consulta: O sistema pode requerer que o usuário faça uma busca explícita pelo especialista ou inferir que competência o usuário deseja encontrar com base nas ações, erros e atividades do mesmo.
5. Operações de interseções (“matching”): refere-se à técnica que será utilizada para fazer a comparação entre os especialistas. Por exemplo: exatidão com a palavra-chave, determinação da similaridade semântica através do método “space-vector” ou uso de algum mecanismo de inferência, por exemplo, se “x” é

um especialista no tópico “y”, talvez também saiba o tópico “z”, ou conheça algum especialista em “z”.

6. Apresentação do resultado: Sistemas de recomendação de especialistas podem apresentar saídas com diferentes graus de detalhes e funcionalidade para posterior exploração. Como no caso dos sistemas de recuperação de informação em que algum mecanismo de ranqueamento é aplicado aos resultados. Entretanto, diferentemente dos sistemas de recuperação de informação, o ranqueamento de especialistas pode ser multidimensional já que um grande número de critérios precisa ser considerado. Além dos detalhes relevantes sobre cada especialista retornado, seu contexto organizacional, contexto social e rede de relacionamentos também podem ser fornecidos. Alguns sistemas também fornecem apenas as fontes em que os indicadores de cada especialista foram identificados, deixando assim a tarefa de escolher o especialista para o usuário.
7. Operações de adaptação e aprendizado: essas operações permitem a customização do resultado retornado através da identificação das preferências do usuário, ou de seu domínio de conhecimento e da coleta de “feedbacks” do usuário a respeito da saída do sistema. O sistema não deve somente identificar o especialista com as competências requeridas, mas também identificar as situações em que o especialista pode ser útil. O sistema pode utilizar o perfil dos usuários para comparar o nível das competências do usuário. Respostas dos usuários referentes à precisão do sistema e ao ranqueamento do resultado podem ser coletados e utilizados pelo sistema para aprender com o tempo. O usuário deve poder adicionar novas competências ou especialistas ao sistema e até mesmo alterar a forma como o resultado do sistema é retornado.

Tabela 15 – Modelo de Domínio Intuitivo para Comparar Sistemas de Busca de Especialista. (YTMAM-SEID & KOBSA, 2003)

Ferramentas	Explicita	Implicita	Dependente de Domínio	Independente de Domínio	Durante a Consulta	Baseada em Agentes	Associada a Modelos Agregados	Busca Explícita	Informação Induzida	Exatidão	Similaridade	Inferência	Lista de Nomes Ordenada	Lista de Nomes Ordenadas e com Informações Pessoais	Especialistas em sua Rede Social	Documentos/Grupo de Especialistas Relevantes	Adaptação usando o Perfil do Especialista	Escolha da Técnica a ser Empregada	Coleta de <i>Feedbacks</i> Relevantes	Usando a Avaliação de Outros
FBEE	X			X			X	X		X				X		X	X			
Expert Finder		X		X	X			X			X			X					X	
P@NOPTIC				X	X			X			X			X						
CNDS		X		X	X				X		X		X							
SAPIENS	X		X					X		X			X							
NAVIGATOR		X		X		X		X	X		X		X				X			
E-QF		X		X			X	X			X			X		X				

Na FBEE o reconhecimento de um especialista é realizado de forma explícita. O pesquisador declara suas competências diretamente no GCC e também pode selecioná-las no conjunto de competências extraídas do Curriculum Lattes ou ainda mapeadas pela mineração dos textos publicados pelos pesquisadores. As operações para extração de indicadores de competências são independentes do domínio, apesar de todas competências serem inferidas a partir do Lattes ou das publicações científicas do pesquisador, elas também poderiam ser provenientes de outras fontes de dados. A construção do perfil emprega a abordagem que associa o especialista a modelos agregados, pois as competências existentes no perfil do especialista estão classificadas na Ontologia de conhecimentos da árvore CNPq.

Na FBEE a busca por especialistas é realizada de forma explícita, ou seja, o usuário seleciona o pesquisador a ser substituído e a partir da análise das competências deste, os candidatos substitutos são apresentados. A operação de “matching” emprega a técnica de exatidão com a palavra-chave, somente os pesquisadores com pelo menos uma competência equivalente a do substituído e as do contexto serão candidatos. Os resultados são apresentados de forma ordenada, onde os gráficos referentes aos pesquisadores mais similares são apresentados mais à esquerda.

As operações de adaptação e aprendizado na FBEE são apoiadas por técnicas de “Visualização de Informação”. Estas possibilitam ao usuário operar sobre o resultado retornado podendo:

1. Obter, através da navegação e diferença de cor dos retângulos providos pelo *Mapa de Árvores*, uma rápida percepção dos contextos de atuação dos pesquisadores e sua importância nestes. Isso possibilita que o usuário aprenda rapidamente quais contextos serão mais afetados com a saída do pesquisador, priorizando a busca por substitutos nos mesmos.
2. Utilizar as visualizações de teia e área para comparar a similaridade dos pesquisadores com base nas competências selecionadas. O resultado provido pode servir como base para uma nova comparação, sendo assim, o usuário pode acrescentar ou remover competências ao conjunto previamente analisado na visualização anterior.

Com isso, a FBEE supera uma limitação presente em grande parte dos sistemas de busca por especialistas que é a forma estática em que o resultado é exibido (YIMAM-SEID & KOOSA, 2003).

Outra característica da FBEE é o filtro por contexto, ou seja, o contexto de atuação do pesquisador a ser substituído é utilizado para reduzir o volume de informação a ser analisado pelo usuário. Neste caso, somente as competências dos pesquisadores comuns ao contexto da busca e seus respectivos itens são utilizados na comparação de similaridades.

A FBEE também permite que um usuário considere outros fatores (perfil MBTI, Filtro de Aprendizado e Filtro de Interesse) para diferenciar pesquisadores com competências equivalentes. Caso as competências de dois ou mais candidatos apresentem graus similares, esses fatores podem ser utilizados como critério de desempate.

Entretanto, o grande diferencial da FBEE está na utilização de diferentes tipos de visualizações para cada faceta do perfil do pesquisador, onde: o *Mapa de Árvores* exibe o contexto de atuação e a importância do pesquisador; os *Gráficos de Teia e Área* exibem as competências do pesquisador relevantes ao contexto de busca; o *Gráfico de Força* ilustra as características do perfil MBTI e por fim os *Gráficos de Linha* exibem a evolução e conhecimento do pesquisador num determinado interesse e competência respectivamente. Dentre todos sistemas analisado, somente o E-QF priorizou o uso de uma visualização mais rebuscada, mesmo assim, somente para exibir a formação de grupos de usuários.

Após esta análise, encerramos a comparação da FBEE com os demais sistemas recomendação de especialistas. A seguir uma breve conclusão sobre o capítulo é descrita.

4.5. Conclusão

Este capítulo apresentou o ambiente GCC, a abordagem de busca proposta neste trabalho e a ferramenta utilizada para execução dessa abordagem. Além disso, uma comparação da FBEE com outros sistemas de recomendação de especialistas também foi realizada.

A FBEE foi implementada no GCC para validarmos nossa abordagem. Esta ferramenta executa uma busca baseada na substituição através de exemplo, a qual utiliza o conhecimento contextual para limitar o conjunto de atributos considerados na identificação

do substituto. Conseqüentemente, o usuário não utilizará informações irrelevantes durante o processo de substituição, obtendo assim resultados mais precisos.

Adicionalmente, o uso de visualizações durante a análise de dados possibilita a agregação da informação e conseqüentemente amplia a cognição, pois permite ao ser humano reparar em detalhes e detectar padrões, trazendo à tona informações que seriam difíceis de analisar e entender.

O próximo capítulo apresenta o estudo de caso elaborado para verificar o quão adequada essa solução é para a busca por pesquisadores substitutos no ambiente GCC.

5. Estudo de Caso

Este capítulo apresenta o estudo de caso realizado para avaliar a BEE, através de seu protótipo funcional FBEE. Este estudo tem o intuito de comparar, qualitativamente, o comportamento da FBEE frente às ferramentas do GCC que podem ser utilizadas no apoio a busca por pesquisadores substitutos. Devido à diferente natureza das informações dos contextos avaliados neste ambiente, a avaliação da abordagem foi dividida em 3 fases. A primeira fase avalia a busca por pesquisadores no contexto de projetos; a segunda avalia a busca em comunidades e, por último, a terceira avalia a busca na organização, isto é, no ambiente como um todo. Nenhum dos entrevistados avaliou mais de 1 módulo. Nesse estudo, o seguinte vocabulário será adotado: um “pesquisador” é uma pessoa cadastrada no GCC, pois no GCC só existem pessoas do meio científico, e um “participante” é algum membro da amostra, alguém que atuou como avaliador neste estudo.

O estudo foi realizado com alunos de graduação, mestrado e doutorado em Ciência da Computação. Eles receberam um treinamento no ambiente GCC para executar as tarefas solicitadas – relacionadas à busca de um substituto - durante a execução do estudo. A seguir, o mesmo grupo de alunos recebeu um treinamento na FBEE para realizarem as mesmas tarefas. De maneira geral, cada participante tinha como tarefa encontrar um ou mais pesquisadores aptos a suprirem a saída de um pesquisador “X”, primeiramente utilizando as ferramentas de busca do GCC e posteriormente a FBEE. Depois de concluídas as tarefas, eles responderam aos questionários referentes à avaliação da proposta.

Os planos dos experimentos seguiram o modelo definido por BARROS *et al.* (2002). De acordo com esses autores, a realização de um estudo experimental geralmente pode ser dividida em cinco fases: a definição, o planejamento, a execução, a análise e o empacotamento do estudo. A definição do estudo consiste em resumir seus objetivos, seu foco de qualidade e os objetos que serão analisados. O planejamento envolve a descrição do perfil dos participantes, dos instrumentos, do processo de execução e uma avaliação crítica dos problemas que podem ser encontrados ao longo desta execução. A execução consiste na realização do estudo experimental pelos participantes, utilizando os instrumentos e o processo definidos no planejamento. A análise consiste na organização dos resultados gerados pelos participantes durante a execução e a realização de inferências sobre estes resultados. Finalmente, o empacotamento consiste na organização e armazenamento dos

documentos construídos nas etapas anteriores, com o intuito de facilitar a repetição do estudo experimental no futuro.

As etapas de descrição e planejamento do estudo foram realizadas em junho de 2008. A instrumentação do estudo, que incluiu a população de uma base com dados fictícios, consumiu o trabalho de um analista de sistemas, membro da equipe do desenvolvimento do GCC.

Inicialmente, definimos as etapas envolvidas no estudo e apresentamos como estas etapas foram realizadas. Em seguida, apresentamos as observações obtidas durante o estudo. Todos os documentos utilizados se encontram no Anexo I.

5.1. Definição

O contexto global considerado neste trabalho é a ausência de recursos que facilitem a escolha de substitutos no cenário científico brasileiro. Já o contexto local tem como objetivo avaliar a viabilidade da utilização da ferramenta desenvolvida, a FBEE, como ferramental de apoio às atividades de busca por substitutos aptos no GCC. O estudo foi desenvolvido tendo em vista a continuidade do desenvolvimento de pesquisas relacionadas com esta abordagem.

Seguindo as diretrizes definidas por WOHLIN *et al.* (2000) e apresentadas em BARROS *et al.* (2002), podemos ressaltar:

- Objeto de Estudo: a FBEE, protótipo funcional da BEE para busca de especialistas por exemplos. Neste caso, um especialista equivale a um pesquisador.
- Objetivo: identificar a viabilidade da utilização da ferramenta desenvolvida, a FBEE, frente às ferramentas do GCC utilizadas no apoio a busca por pesquisadores substitutos.
- Foco de Qualidade: os ganhos obtidos pela utilização da ferramenta proposta, medidos através da precisão dos resultados retornados e tempo gasto para obtê-los, além das dificuldades e satisfação dos participantes.
- Perspectiva: O estudo será desenvolvido sob a ótica do profissional da Ciência, que executa atividades científicas, avaliando a viabilidade de utilização da

ferramenta proposta, tendo em vista a continuidade do desenvolvimento das pesquisas relacionadas com as técnicas e ferramentas criadas. Sendo assim, tais profissionais avaliarão a eficiência da ferramenta na busca de substitutos e a viabilidade de utilização da mesma.

- Contexto: A demonstração da ferramenta, com uso de dados fictícios, no ambiente GCC.

Seguindo a notação GoalQuestionMetric (GQM) (SOLINGEN & BERGHOUT 1999), a definição do estudo é:

Analisar a viabilidade de utilização da FBEE para apoiar processos de Gestão do Conhecimento referente à uma melhor busca por substitutos.

Com o propósito de recolher requisitos de melhorias e avaliar a viabilidade do seu uso no ambiente GCC.

Referente aos possíveis ganhos obtidos por seu uso e as dificuldades encontradas.

Do ponto de vista do pesquisador.

No contexto de gestão do conhecimento em ambientes científicos, no cenário nacional.

5.2.Planejamento

• Cenário Utilizado

O estudo investigativo contou com a participação de profissionais da área de Computação, os quais são alunos de bacharelado dos cursos de Ciência da Computação da Universidade Federal do Rio de Janeiro, além de alunos de mestrado e doutorado da COPPE, do Programa de Engenharia de Sistemas e Computação (PESC).

O estudo não foi executado em um ambiente comercial ou industrial, mas em um ambiente acadêmico. A capacidade de generalização deste estudo é discutida adiante, quando avaliamos as limitações e problemas que foram encontrados durante sua execução.

• Participantes

Os participantes do estudo investigativo foram divididos em três grupos de 4 pessoas, sendo que nenhum participante integrou mais de um grupo. O primeiro grupo foi formado por 2 alunos de graduação e 2 de doutorado. O segundo por 2 alunos de graduação, 1 de mestrado e 1 de doutorado. Finalmente, o terceiro por 2 alunos de

graduação e 2 de mestrado. Por causa da indisponibilidade dos participantes, não foi possível contrabalançar os grupos.

A opção por grupos distintos foi feita com o intuito de manter certa heterogeneidade entre os participantes das amostras em questão, já que a viabilidade da ferramenta foi testada em 3 contextos diferentes (projeto, comunidade e a organização). Para cada um dos grupos foi feito um estudo de caso separado.

- **Treinamento**

Todos receberam treinamento formal de como executar uma busca por substitutos através do GCC e, posteriormente, de como realizar a mesma busca através da FBEE. Além disto, o desenvolvedor da BEE e FBEE estava à disposição para monitoração do uso e suporte. Cada grupo recebeu seu treinamento separadamente.

- **Instrumentos**

Foi disponibilizado para cada participante um computador para acessar a FBEE e o ambiente GCC. A base de dados utilizada possui ao todo 192 pesquisadores cadastrados, os quais tiveram grande parte de suas informações mineradas dos currículos Lattes de pesquisadores de instituições avaliadas pela CAPES como nível 5, 6 ou 7 (CAPES, 2008). Destes pesquisadores, 113 são do PESC e 79 de outras instituições. Além disso, essa base ainda contém 15 projetos e 35 comunidades cadastradas. Grande parte destes pesquisadores pertence a um ou mais projetos e comunidade, do nosso contexto de busca. Entretanto, devido ao baixo número de competências similares entre os pesquisadores existentes e entre pesquisadores e contextos de atuação, foi preciso atribuir novas competências a alguns desses pesquisadores e seus respectivos contextos para que a ferramenta pudesse ser avaliada de maneira mais precisa.

Todos participantes receberam um questionário de caracterização (Q1), para levantamento de sua formação e experiência, e um questionário (Q5) para avaliação da satisfação com a ferramenta BEE. Os questionários para comparação da FBEE com as ferramentas de busca do GCC nos contextos de projeto, de comunidade e organizacional (Q2, Q3 e Q4) foram entregues, respectivamente, aos participantes do Grupo I, Grupo II e Grupo III. Conforme dito anteriormente, esses questionários podem ser encontrados no Anexo I.

- **Crítérios**

O foco de qualidade do estudo exige critérios -- tais como usabilidade da ferramenta, satisfação com os resultados obtidos, tempo gasto na busca e coerência dos resultados encontrados por diferentes participantes -- que avaliem os possíveis ganhos obtidos e as dificuldades encontradas na FBEE. Tanto os ganhos quanto as dificuldades foram avaliados qualitativamente, através dos questionários já mencionados. Esta análise tem o objetivo de avaliar a ferramenta, as dificuldades e possibilidades de melhorias futuras.

- **Hipótese Nula**

A hipótese nula é uma afirmativa que o estudo investigativo tem como objetivo negar. No estudo atual, a hipótese nula determina que a utilização da FBEE NÃO produz benefícios no que se refere à busca por substitutos aptos no GCC. De acordo com os critérios selecionados, esta hipótese se traduz na inexistência de diferenças significativas utilizando tal solução em relação às ferramentas convencionais do GCC.

- **Hipótese Alternativa**

A hipótese alternativa é uma afirmativa que nega a hipótese nula. O estudo experimental tem como objetivo provar a hipótese alternativa, refutando assim a hipótese nula. No estudo atual, a hipótese alternativa determina que os participantes do estudo que avaliaram a FBEE podem ter benefícios no que se refere à busca por substitutos no GCC.

- **Variáveis Independentes**

As áreas de atuação, os dados pessoais, experiência em liderança e a escolaridade dos participantes são informações independentes coletadas durante o estudo.

- **Variáveis Dependentes**

Todas as demais variáveis (utilidade, interesse, funcionalidades, vantagens, viabilidade, agrado, desagrado, sugestão de funcionalidades e observação) são dependentes.

- **Análise Qualitativa**

Tem o objetivo de avaliar o que mais agradou e o que mais desagradou na proposta, além de dificuldades e sugestões.

- **Capacidade Aleatória**

Pode ser exercida na seleção dos participantes do estudo e na distribuição dos objetos de análise entre eles. Infelizmente, por questão de disponibilidade, a seleção

aleatória não foi possível, mas os objetos de análise foram distribuídos aleatoriamente entre os participantes.

- **Classificação em Bloco**

Devido à diferente natureza das informações dos contextos avaliados, houve necessidade de se formar 3 blocos para analisarem a solução para substitutos em projeto, comunidade e contexto organizacional.

- **Balanceamento**

Apesar dos grupos possuírem a mesma quantidade de participantes, essa distribuição foi ocasional e não planejada. Entretanto, por questão de indisponibilidade, não conseguimos homogeneizar os grupos a partir do grau de escolaridade dos participantes. Certamente, grupos mais homogêneos – com 1 aluno de doutorado, 1 de mestrado e dois de graduação – aumentariam a confiabilidade dos resultados obtidos. Através destes, poderíamos averiguar, com maior precisão, se a formação dos participantes influencia os resultados obtidos, ou seja, se os participantes de cada grupo – independentemente da sua formação – recomendariam o mesmo substituto.

- **Mecanismo de Análise**

As variáveis dependentes são apresentadas utilizando-se as escalas próprias de cada variável. Além disto, os resultados são discutidos utilizando-se por base análise baseada no ranqueamento dos resultados obtidos pelos participantes.

- **Validade Interna do Estudo**

A validade interna de um estudo é definida como a capacidade de um novo estudo repetir o comportamento do estudo atual com os mesmos participantes e objetos com que ele foi realizado. A validade interna do estudo é dependente do número de participantes executando o estudo. Infelizmente somente 12 pessoas participaram do nosso estudo, o que resultou na formação de 3 grupos de 4 pessoas para validar a substituição em cada um dos 3 contextos de busca. Como o objetivo central do estudo é que os participantes indiquem o mesmo substituto em cada um dos contextos, certamente, um número maior de participantes melhoraria esta validação.

Outro ponto que pode influenciar o resultado do estudo é a troca de informações entre os participantes que já realizaram o estudo e os que não o realizaram. Para evitar este

problema, requisitamos explicitamente que os participantes não trocassem informações a respeito da ferramenta e nem se comunicassem enquanto preenchiam os questionários.

- **Validade Externa do Estudo**

A validade externa do estudo mede sua capacidade de refletir o mesmo comportamento em outros grupos de participantes e profissionais do meio acadêmico, ou seja, em outros grupos além daquele em que o estudo foi aplicado. A validade interna do estudo foi considerada suficiente, visto que o presente estudo visa avaliar a viabilidade de uma ferramenta. Demonstrada esta viabilidade, novos estudos podem ser planejados para refinar a solução.

- **Validade de Construção do Estudo**

A validade de construção do estudo se refere à relação entre os instrumentos e participantes do estudo e a teoria que está sendo provada por este. Neste caso, todos os participantes são do meio científico, os dados utilizados para demonstração são fictícios e dizem respeito a projetos, comunidades e pesquisadores. Além disso, o estudo não visa avaliar a correção do cenário, mas a viabilidade do uso da FBEE.

- **Validade de Conclusão do Estudo**

A validade de conclusão do estudo mede a relação entre os tratamentos e os resultados, determinando a capacidade do estudo em gerar alguma conclusão. Não encontramos grandes dificuldades em relação à capacidade de conclusão do estudo, visto que os seus resultados são quantitativos.

5.3.Execução

De maneira geral, cada participante tinha como trabalho encontrar um substituto “mais” apto a suprir a saída de um pesquisador “X”, primeiramente utilizando as ferramentas de busca do GCC e posteriormente utilizando a FBEE. O seguinte problema foi apresentado aos participantes: “Suponha que o Pesquisador 1(Pesq1) vai abandonar a instituição em questão, quais seriam os substitutos indicados para substituí-lo?”. A busca ocorreu em 3 contextos diferentes:

1. Projeto: A substituição ocorreu no “Projeto 1” (Proj1), composto por 23 pesquisadores, (Pesq1) era gerente em (Proj1). Resultado esperado: Pesquisador 4 (Pesq4).

2. Comunidade: A substituição ocorreu na “Comunidade 1” (Com1), composta por 32 pesquisadores, (Pesq1) era moderador em (Com1). Resultado esperado: Pesquisador 8 (Pesq8).
3. Organizacional: A substituição ocorreu na organização em questão, composta por 50 pesquisadores. Repare que apesar de existirem 192 pesquisadores na base, somente os que realizaram algum tipo de interação no ambiente e tiveram parte de suas informações alteradas foram considerados. Resultado esperado: Pesquisador 8 (Pesq8).

O estudo investigativo foi dividido em duas etapas. Na primeira parte, os participantes receberam um treinamento no ambiente GCC para executar a substituição solicitada ao longo da execução do estudo. Na segunda parte, o mesmo grupo recebeu um treinamento na FBEE para executar a mesma tarefa. Os participantes foram distribuídos em três grupos, os quais realizaram o experimento em momentos distintos. O Grupo I avaliou a substituição no contexto de projetos e respondeu os questionários Q1, Q2 e Q5. O Grupo II avaliou a substituição no contexto de comunidades e respondeu os questionários Q1, Q3 e Q5. Por fim, o Grupo III avaliou a substituição no contexto organizacional e respondeu os questionários Q1, Q4 e Q5.

Os questionários foram distribuídos no final de cada etapa e recolhidos antes do início da próxima, sendo que ambas as etapas ocorreram no mesmo dia. Durante o tempo de resposta dos questionários não houve comunicação entre os participantes.

5.4. Análise dos Resultados

Nesta seção são apresentados os resultados do estudo investigativo da FBEE. Conforme dito, 3 grupos participaram desse estudo, sendo assim, sempre que necessários os resultados de cada grupo são exibidos separadamente. Outra informação importante refere-se aos questionários respondidos por cada grupo, ou seja: Grupo I (Q1, Q2 e Q5); Grupo II (Q1, Q3 e Q5); e Grupo III (Q1, Q4 e Q5).

5.4.1. Perfil dos Participantes

Dentre os questionários preenchidos pelos participantes, o primeiro (Q1) foi utilizado para extrair o perfil de cada um. Todos são da área de computação, sendo 3 gerentes de projetos, 9 analistas e exercem tais funções na COPPE/UFRJ. Todos os gerentes possuem uma equipe pequena, composta no máximo por 10 pessoas. O ideal seria ter um gerente em cada grupo, porém por questões de incompatibilidade de horários, 2 gerentes tiveram que participar do mesmo grupo. Essas informações são ilustradas na Tabela 16.

Tabela 16 – Informações detalhadas sobre os participantes do estudo.

ID	Grupo	Formação	Exp. Desenvolvimento	Gerente (Tam. Equipe)
1	I	D.Sc.	Acadêmica (8 anos)	Sim(0-10)
2	I	D.Sc.	Acadêmica (6 anos)	Sim(0-10)
3	I	Graduando	Acadêmica (2 anos)	Não
4	I	Graduando	Acadêmica (2 anos)	Não
5	II	D.Sc.	Acadêmica (6 anos)	Sim(0-10)
6	II	M.Sc	Acadêmica (4 anos)	Não
7	II	Graduando	Acadêmica (1 ano)	Não
8	II	Graduando	Acadêmica (1 ano)	Não
9	III	M.Sc	Acadêmica (3 anos)	Não
10	III	M.Sc	Acadêmica (3 anos)	Não
11	III	Graduando	Acadêmica (3 anos)	Não
12	III	Graduando	Acadêmica (3 anos)	Não

5.4.2. Análise Comparativa da FBEE e das Ferramentas de Busca do GCC

Um dos objetivos deste estudo é conhecer a percepção dos participantes sobre a FBEE frente às funcionalidades de busca do GCC. As informações referentes à análise comparativa estão presentes nos questionários Q2, Q3 e Q4, respondidos pelos Grupos I, II e III, respectivamente.

5.4.2.1. Avaliação da Estrutura e Organização do Trabalho

A Tabela 17 mostra os resultados referentes à organização e estruturação do trabalho. A primeira coluna apresenta as perguntas feitas aos participantes. As quatro colunas seguintes mostram as respostas dos participantes, onde S = Sim; N = Não; R = Razoável e NA = Não se aplica. De maneira geral, todos os participantes ficaram satisfeitos com a organização e estruturação do trabalho, somente 1(um) participante (ID 5) do Grupo II a considerou “razoável”, porém não justificou o motivo.

Tabela 17 – Satisfação quanto a apresentação e estruturação do trabalho.

Sobre a organização e estrutura do trabalho				
	S	N	R	NA
1. Você ficou satisfeito com a apresentação do trabalho?	11		1	
2. Você ficou satisfeito com o treinamento referente ao ambiente GCC?	11		1	
3. Você ficou satisfeito com o treinamento referente a FBEE?	11		1	
4. Você ficou satisfeito com o treinamento com um todo?	11		1	

5.4.2.2. Busca por Substitutos Através do GCC

Os tópicos desta seção referem-se à avaliação da busca por substitutos através das ferramentas disponibilizadas no GCC.

Critérios considerados na busca de um substituto

Como os critérios de busca foram definidos por cada participante separadamente, já que estes mudam de acordo com o contexto da busca, as Tabelas 18, 19 e 20 foram criadas para cada grupo com base nas respostas de seus respectivos participantes. A primeira linha referencia o “Grupo” entrevistado; a primeira coluna contém os critérios utilizados; por fim, as quatro colunas a seguir apresentam os participantes e os critérios utilizados pelos mesmos.

Tabela 18 – Itens utilizados na busca por substitutos no contexto de projetos.

Grupo I				
	ID1	ID2	ID3	ID4
Competências particulares	X	X	X	X
Ser gerente de projeto	X	X		
Disponibilidade para trabalhar	X			

Tabela 19 – Itens utilizados na busca por substitutos no contexto de comunidades.

Grupo II				
	ID5	ID6	ID7	ID8
Competências particulares	X	X	X	X
Competências da comunidade	X	X		
Competências do projeto	X	X		

Tabela 20 – Itens utilizados na busca por substitutos no contexto organizacional.

Grupo III				
	ID9	ID10	ID11	ID12
Competências particulares	X	X	X	X
Competências da comunidade	X			
Competências do projeto	X	X		
Perfil MBTI	X			

Conforme podemos observar nas três tabelas acima, todos os participantes, independente do contexto, consideraram as competências particulares dos pesquisadores na busca por substitutos. Isto é bastante intuitivo, pois sempre que procuramos um substituto, buscamos alguém que apresente os mesmos conhecimentos da pessoa a ser substituída. Entretanto, poucos participantes consideraram a relevância da competência no contexto de busca, que na nossa opinião é um critério importante.

Outros itens como perfil MBTI, importância do substituto no contexto e a disponibilidade para trabalhar foram pouco utilizados, mostrando que os participantes consideram as competências, por si só, um critério suficiente ao processo de substituição.

Um ponto curioso em relação ao Grupo I, conforme mostrado na Tabela 18, é que os dois participantes (ID1 e ID2) que consideraram o item *ser gerente de projeto* são gerentes e alunos de doutorado. Isso mostra que pelo fato de serem gerentes, eles sabem o quanto à experiência gerencial pode ser relevante ao se recomendar um substituto.

Outra curiosidade refere-se ao Grupo II, vide Tabela 19, no qual dois de seus participantes (ID5 e ID6) consideraram as competências do projeto, mesmo realizando suas buscas dentro da “Comunidade 1”. Isto mostra que, para estes participantes, a experiência relacionada às atividades profissionais é um critério importante numa comunidade.

Funcionalidades utilizadas na busca de um substituto

Os participantes receberam treinamento nas 6(seis) funcionalidades do GCC que poderiam ser utilizadas na busca por um substituto, são elas: Busca por Projetos, Busca por Comunidades, Busca por Participantes, Busca por Competências, Busca por Pesquisadores no módulo de Inteligência Competitiva e Ferramentas de Análises.

A *Busca por Projetos* permite ao usuário obter todos os projetos relacionados a uma palavra-chave, pesquisador, competência e área de conhecimento. A *Busca por Comunidades* permite obter todas as comunidades relacionadas a uma palavra-chave, pesquisador, competência e área de conhecimento. A *Busca por Participantes* permite buscar um pesquisador através do nome, competências, áreas de conhecimento e perfil MBTI. A *Busca por Competências* possibilita buscar pesquisadores através de suas competências, a diferença desta busca em relação as anteriores é que a relevância do pesquisador em cada competência é apresentada. A *Busca por Pesquisadores no módulo de Inteligência Competitiva* é similar a anterior, porém esta só utiliza as competências mineradas no Lattes do pesquisador, enquanto a *Busca por Competências* utiliza as competências mineradas no Lattes e as cadastradas no GCC manualmente. Por último, as *Ferramentas de Análise* que disponibilizam gráficos de “pizza” ou “barra” referentes aos itens relacionadas a projetos, comunidades e pesquisadores cadastrados no sistema.

A Tabela 21 mostra os resultados referentes a utilização dessas 6(seis) funcionalidades de busca. A primeira e segunda coluna apresentam respectivamente o ID e Grupo dos Participantes. As cinco colunas seguintes mostram as funcionalidades e suas respectivas utilizações, onde BP = Busca por Projetos; BC = Busca por Comunidades; BPA = Busca por Participantes; BCO = Busca por Competências; BPE = Busca por Pesquisador e FA = Ferramentas de Análise.

Tabela 21 – Funcionalidades utilizadas na busca por pesquisadores substitutos

ID	Grupo	BP	BC	BPA	BCO	BPE	FA
1	I	X			X		
2	I	X			X		X
3	I				X		
4	I				X		
5	II	X	X	X			
6	II	X	X				
7	II			X	X		
8	II				X		
9	III	X	X		X		
10	III		X		X		
11	III			X			
12	III			X			
Total Gupo I		2	0	0	4	0	1
Total Gupo II		2	2	2	2	0	0
Total Gupo III		1	2	2	2	0	0
Total		5	4	4	8	0	1

De um modo geral, os participantes preferiram a BCO pelo fato de exibir a relevância do pesquisador nos conjuntos de competências utilizadas. A seguir, a BP foi a funcionalidade mais importante. No Grupo I está funcionalidade foi utilizada para identificar a disponibilidade do pesquisador e sua experiência gerencial, enquanto nos Grupos II e III a BP foi utilizada na identificação dos projetos relacionados às competências do pesquisador a ser substituído (Pesq1). Outras funcionalidades com utilização razoável

foram a BC e a BPA. A BC foi utilizada na identificação dos pesquisadores pertencentes a “comunidade 1”, isto é, da qual (Pesq1) estava saindo. A BPA foi utilizada na identificação dos pesquisadores que possuíam competências similares a (Pesq1).

As FA foram utilizadas somente por um participante (ID2), o que mostra uma rejeição inicial dos participantes às ferramentas gráficas. A BPE não foi utilizada pelo fato da BCO englobar, conforme dito anteriormente, um maior conjunto de competências no seu escopo de busca.

Tempo gasto na busca por substitutos

A Tabela 22 ilustra o tempo gasto por cada participante na busca de um substituto através do GCC. A primeira e segunda coluna apresentam respectivamente o ID e Grupo do participante, enquanto a terceira apresenta o tempo gasto por cada pessoa.

Tabela 22 – Tempo gasto na busca por substitutos através do GCC.

ID	Grupo	Tempo(min)
1	I	30
2	I	46
3	I	34
4	I	35
5	II	21
6	II	25
7	II	32
8	II	30
9	III	32
10	III	18
11	III	23
12	III	26

O tempo médio da busca foi de 29.3 minutos. No geral, todos participantes – independentes da sua formação – ficaram próximos dessa média, exceção feita a dois participantes (ID2 e ID10). O primeiro (ID2) é um dos gerentes de projeto, sua demora foi decorrente da insatisfação com os resultados obtidos, visto que foi o único a utilizar ferramentas de análise na tentativa de melhorar as respostas, conforme mostrado na Tabela

21. Já o outro participante (ID10) - aluno de Mestrado - foi o que levou menos tempo para encontrar um substituto. Isto foi decorrente da pouca paciência que ele teve para analisar os resultados a partir do GCC.

Possíveis substitutos encontrados

A Tabela 23 mostra o resultado da busca. O ID e Grupo de cada participante estão representados na primeira e segunda coluna respectivamente. A terceira contém o contexto em que a busca foi realizada. Finalmente, a última coluna apresenta os possíveis substitutos, ordenados por relevância. Devemos lembrar que os critérios considerados variam de acordo com o contexto, logo os resultados devem ser analisados dentro de cada grupo isoladamente. Lembrando sempre que (Pesq1) é o pesquisador a ser substituído em todos os contextos.

Tabela 23 – Substitutos recomendados.

ID	Grupo	Contexto	Substituto
1	I	Projeto 1	Pesq 10, Pesq 5
2	I		Pesq 4, Pesq 2
3	I		Pesq 12, Pesq 8, Pesq 3
4	I		Pesq 2, Pesq 8
5	II	Comunidade 1	Pesq 2, Pesq 3, Pesq 4, Pesq 6
6	II		Pesq 2, Pesq 4, Pesq 8
7	II		Pesq 8, Pesq 4
8	II		Pesq 3, Pesq 2
9	III	Organizacional	Pesq 2, Pesq 4
10	III		Pesq 8, Pesq 9
11	III		Pesq 11
12	III		Pesq 4

Analisando os resultados acima, observamos a discrepância de indicações existentes dentro de cada grupo. Somente no Grupo II obtivemos uma recomendação mais homogênea (Pesq 2). Essa discrepância foi decorrente dos critérios que cada um considerou durante a execução da busca por substitutos, mostrando que a decisão é uma coisa muito pessoal.

Logo, é recomendado que os executores sempre utilizem um conjunto de critérios comuns durante o processo substituição.

Quanto à precisão dos resultados, somente um participante de cada grupo chegou ao resultado esperado – (Pesq4) no Grupo I e (Pesq8) nos Grupos II e GIII – o que representa uma precisão 25%. Valor considerado que consideramos baixo.

Outro fator que chamou atenção refere-se à insatisfação dos participantes quanto aos resultados obtidos. Os comentários são sumarizados a seguir:

- A necessidade de analisar informações em locais diferentes prejudica o processo de escolha.
- Dificuldades em comparar os pesquisadores com base em suas competências.
- A necessidade de dar muitos “cliques” para chegar à informação procurada é um problema.
- A análise de cada substituto individualmente é um processo árduo.
- Falta um mecanismo mais preciso e amigável para comparar os possíveis substitutos.
- Muitos dados a serem considerados, tornando a busca um processo cansativo.
- Não consegui analisar a rede social dos pesquisadores.

Durante o processo de busca através do GCC, a maior parte dos participantes utilizou papel e caneta ou mais de uma tela do sistema para facilitar a análise dos relatórios retornados. O ideal seria o participante fazer toda essa análise através de uma única tela e sem auxílio de objetos externos.

A próxima seção apresenta a análise dos resultados obtidos com a utilização da FBEE, os quais são utilizados em comparação aos resultados obtidos nesta seção.

5.4.3. Busca por Substitutos através da FBEE

Os tópicos desta seção referem-se à avaliação da busca por substitutos através da FBEE.

Critérios considerados na busca de um substituto

Diferentemente do que ocorreu durante a utilização do GCC, os critérios utilizados foram definidos previamente para cada contexto, devendo o candidato valorar (1-5) os que considera importante. O objetivo deste tópico é identificar se os participantes utilizariam os

mesmos critérios de busca, a partir do momento que critérios não considerados fossem explicitados a eles.

As Tabelas 24, 25 e 26 foram criadas separadamente para cada grupo. A primeira linha referencia o “Grupo” entrevistado. A primeira coluna contém os critérios utilizados, enquanto as quatro colunas a seguir apresentam os participantes e suas respectivas respostas, finalmente a última coluna apresenta a média dos resultados para um total de 20 pontos.

Tabela 24 – Critérios utilizados na busca de por substitutos em projetos.

Grupo I					
	ID1	ID2	ID3	ID4	Média
Interatividade no projeto	3	1	1	1	30%
Ser gerente de projeto	3	4	1	1	45%
Menor discrepâncias entre as competências requeridas pelo projeto e pelos possíveis substitutos.	4	5	3	5	85%
Menor discrepâncias entre as competências requeridas pelo pesquisador a ser substituído e seus possíveis substitutos.	4	5	4	4	85%
Grau de conhecimento na competência do pesquisador a ser substituído.	4	5	5	4	90%
Grau de interesse na competência do pesquisador a ser substituído.	4	1	1	1	35%
Similaridade entre perfil (MBTI)	2	1	1	1	25%

Analisando o resultado do Grupo I, conforme mostrado na Tabela 24, observamos que os participantes mantiveram-se fiéis aos critérios escolhidos previamente na busca através do GCC, vide Tabela 18, porém adicionaram outros 2(dois):

- *Grau de conhecimento na competência do pesquisador a ser substituído*: neste caso, todos participantes perceberam que a mensuração da competência pode ser um diferencial na escolha do possível substituto.
- *Menor discrepâncias entre as competências requeridas pelo projeto e pelos possíveis substitutos*: todos participantes concordaram que as competências

relevantes ao contexto sempre devem ser consideradas durante a busca. Um dos participantes (ID2) fez o seguinte comentário: “Fiquei tão preocupado em buscar um substituto com competências similares, que acabei me esquecendo de considerar quais competências da pessoa a ser substituído são realmente relevantes ao *projeto I*”.

Tabela 25 – Critérios utilizados na busca de por substitutos em comunidades.

Grupo II					
	ID5	ID6	ID7	ID8	Média
Interatividade na comunidade	4	3	3	4	70%
Ser moderador de comunidades	4	1	2	3	50%
Ser ponte entre comunidades	4	1	2	1	40%
Menor discrepâncias entre as competências requeridas pela comunidade e pelos possíveis substitutos.	5	4	5	4	90%
Menor discrepâncias entre as competências requeridas pelo pesquisador a ser substituído e seus possíveis substitutos.	4	5	4	5	90%
Grau de conhecimento no assunto do pesquisador a ser substituído.	1	1	4	4	50%
Grau de interesse no assunto do pesquisador a ser substituído.	1	1	3	2	35%
Similaridade entre perfil (MBTI)	1	4	1	1	35%

Analisando o resultado do Grupo II, exibidos na Tabela 25, e comparando-o com o resultado da Tabela 19, alguns critérios chamaram atenção:

- *Menor discrepâncias entre as competências requeridas pela comunidade e pelos possíveis substitutos*: do mesmo modo que em projetos, todos participantes concordaram que esse critério é fundamental na busca do substituto. Anteriormente somente 2 participantes (ID5 e ID6) deste grupo observaram essa importância.

- *Interatividade na comunidade*: anteriormente desconsiderado, este critério passou a ter uma boa aceitação. Os participantes chegaram à conclusão que um bom substituto neste contexto deve atuar ativamente na comunidade.
- *Menor discrepâncias entre as competências requeridas pelo pesquisador a ser substituído e seus possíveis substitutos*: este critério continuou sendo o mais importante para os participantes, conforme esperávamos.

Tabela 26 – Critérios utilizados na busca de por substitutos na organização.

Grupo III					
	ID9	ID10	ID11	ID12	Média
Participação em comunidades	5	4	2	3	70%
Interatividade na comunidade	4	4	3	2	65%
Ser moderador de comunidades	5	2	3	3	65%
Ser ponte entre comunidades	2	2	3	1	40%
Participação em projetos	5	3	1	3	60%
Interatividade em projetos	2	3	2	2	45%
Ser gerente de projetos	3	2	2	3	50%
Menor discrepâncias entre as competências requeridas pelo substituto e o pesquisador a ser substituído.	4	5	5	5	95%
Grau de conhecimento no assunto do pesquisador a ser substituído.	4	5	5	4	90%
Grau de interesse no assunto do pesquisador a ser substituído.	2	2	5	3	60%
Similaridade entre perfil (MBTI)	5	1	5	2	65%

Comparando o resultado do Grupo III, conforme mostrados na Tabela 26, em relação ao do GCC, mostrado na Tabela 20, observamos que muitos critérios previamente desconsiderados ou pouco utilizados na busca realizada através das ferramentas do GCC receberam boas notas. Este fato demonstra que numa busca organizacional, além das competências particulares, a importância e a interatividade dos possíveis substitutos em seus contextos de atuação são extremamente relevantes.

Um critério que gerou grande debate foi a *similaridade de perfil MBTI*. A maior parte dos participantes não utilizaram este critério por considerá-lo irrelevante. Um dos participantes (ID1) teceu o seguinte comentário: “Os psicólogos trabalham meses para encontrar perfis similares através do teste MBTI, acho pretensioso demais você afirmar que as informações existentes no sistema são suficientes para detectar tal similaridade”.

Tempo gasto na busca por substitutos

A Tabela 27 ilustra o tempo gasto por cada participante na busca de um substituto através da FBEE. A primeira e segunda coluna apresentam respectivamente o ID e Grupo do Participante, enquanto a terceira apresenta o tempo gasto por cada um.

Tabela 27 – Tempo gasto na busca por substitutos através da FBEE.

ID	Grupo	Tempo(min)
1	I	15
2	I	20
3	I	19
4	I	22
5	II	15
6	II	20
7	II	18
8	II	17
9	III	20
10	III	12
11	III	8
12	III	16

Analisando os resultados da Tabela 27 em relação aos da Tabela 22, observamos que todos os participantes realizaram a busca por substitutos num menor tempo, em alguns casos a diferença chegou a aproximadamente 50% (ID1, ID2 e ID11). O tempo médio de busca através da FBEE foi de 15.5 minutos, o que representa uma redução de aproximadamente 47% em relação ao tempo médio anterior (29.3 minutos). Essa redução é

um fator extremamente importante, visto que muitos dos participantes haviam reclamado do tempo necessário para encontrar e analisar as informações relevantes através do GCC.

Possíveis substitutos encontrados

A Tabela 28 mostra o resultado da busca por substitutos. As duas primeiras colunas apresentam respectivamente o ID e Grupo do participante. Enquanto a terceira contém o contexto em que a busca foi realizada. Finalmente, a última coluna apresenta os possíveis substitutos, ordenados por relevância.

Tabela 28 – Substitutos recomendados através da FBEE.

ID	Grupo	Contexto	Substituto
1	I	Projeto 1	Pesq 5, Pesq 4
2	I		Pesq 4, Pesq 3
3	I		Pesq 4, Pesq 3
4	I		Pesq 4
5	II	Comunidade 1	Pesq 4, Pesq 11, Pesq 8
6	II		Pesq 8, Pesq 2
7	II		Pesq 8
8	II		Pesq 8, Pesq 4
9	III	Organizacional	Pesq 8, Pesq 11, Pesq 4
10	III		Pesq 8, Pesq 5
11	III		Pesq 8, Pesq 3, Pesq 11
12	III		Pesq 8, Pesq 3

A partir desta tabela, podemos observar que os participantes de cada grupo obtiveram resultados bastante similares, mostrando que um conjunto de critérios comum é fundamental no processo de busca por um substituto:

No Grupo I, o (Pesq 4) foi apontado como possível substituto por todos os participantes, sendo que 75% o apontaram como o “mais” apto a substituir (Pesq 1), que era o substituto esperado. Somente um dos participantes (ID1) não considerou o (Pesq 4) como mais relevante. Essa disparidade foi decorrente dos critérios utilizados na avaliação: Enquanto os demais

participantes concentraram suas escolhas na similaridade e grau das competências em relação ao pesquisador a ser substituído e ao contexto, o (ID1) considerou também a disponibilidade do possível substituto.

No Grupo II, o (Pesq 8) – substituto que considerávamos ideal – foi apontado como possível substituto por todos os participantes, onde 75% o apontaram como o mais apto a substituir (Pesq 1) na *Comunidade 1*. Apenas um dos participantes (ID5) não considerou o (Pesq 8) como mais relevante. Este fato também foi decorrente dos critérios utilizados por (ID5): Enquanto os outros participantes concentraram suas escolhas no grau das competências similares as do pesquisador a ser substituído e as relevantes ao contexto, o ID5 considerou somente as competências relevantes a *Comunidade 1*.

No Grupo III, (Pesq 8) foi apontado, por todos os participantes, como o substituto mais apto a substituir (Pesq 1) na *Organização*. Vale lembrar que (Pesq 8) era considerado o substituto ideal neste grupo. Este resultado foi consequência de todos os participantes terem considerado, principalmente, a discrepância da competência durante a comparação entre o pesquisador a ser substituído e seus possíveis substitutos. Uma curiosidade é que ID5 utilizou a similaridade de perfil MBTI para decidir a prioridade dos substitutos por ele indicado, e coincidentemente (Pesq 8) apresentou o perfil mais próximo a (Pesq 1).

Quando comparamos os resultados obtidos através do GCC, ilustrados na Tabela 23, com os obtidos pela FBEE, conforme mostrados na Tabela 28, observamos que somente dois dos participantes (ID2) e (ID10) encontraram o mesmo substituto prioritário (Pesq 4) e (Pesq 11), respectivamente. No caso de (ID2) acreditamos que essa similaridade está relacionada ao fato dele ter sido o participante que levou mais tempo (46 minutos) durante a análise através do GCC, conforme mostrado na Tabela 22. Em relação ao (ID10) acreditamos que tal similaridade está relacionada ao fato dele ter considerado os mesmos critérios durante ambas as buscas (GCC e FBEE).

Apesar da precisão dos resultados obtidos através da FBEE serem bem superiores aos dos obtidos através das ferramentas do GCC. As conclusões do parágrafo anterior

sugerem a realização de um estudo mais detalhado, ou seja, com um maior número de participantes e com grupos mais homogêneos.

Avaliação das visualizações oferecidas

A Tabela 29 apresenta os resultados obtidos na avaliação das visualizações oferecidas pela FBEE. A primeira coluna apresenta o ID dos participantes, enquanto as cinco seguintes mostram as visualizações utilizadas, onde MA = Mapa de Árvores; GTA = Gráfico de Teia e Área; GLI = Gráfico de Linha para mensurar Interesses; GLC = Gráfico de Linha para mensurar Competências e GF = Gráfico de Força para similaridade de perfil MBTI. Todas as visualizações utilizadas receberam nota de 1-5, caso contrário receberam NA = Não se Aplica. A penúltima linha da tabela apresenta a média das notas atribuídas pelos participantes a cada uma das visualizações, as atribuições NA foram desconsideradas no cálculo desta média.

Tabela 29 – Avaliação das visualizações providas pela FBEE.

ID	MA	GTA	GLI	GLC	GF
1	5	5	3	3	2
2	5	5	NA	NA	NA
3	5	5	NA	NA	NA
4	5	5	NA	NA	NA
5	5	5	4	4	NA
6	5	5	4	4	4
7	5	5	4	4	3
8	4	4	NA	NA	NA
9	5	4	5	5	5
10	5	5	NA	NA	4
11	4	5	NA	NA	5
12	4	4	NA	NA	4
Média Grupo I	5.0	5	3.0	3.0	2.0
Média Grupo II	4.75	4.75	4.0	4.0	3.75
Média Grupo III	4.5	4.5	5.0	5.0	4.5
Média	4.75	4.75	4.0	4.0	3.85

A visualização MA (Mapa de Árvores), usada para avaliar a importância e interatividade do pesquisador em seu contexto de atuação, obteve grande aceitação por parte dos participantes. Os principais comentários são sumarizados a seguir:

- Os resultados são apresentados de forma clara e objetiva.
- O resumo referente a cada pesquisador ao se posicionar o cursor do “mouse” sobre seu respectivo retângulo é de grande utilidade.
- A visualização do todo facilita a identificação da importância e de possíveis substitutos em cada contexto.

Um dos participantes (ID11) teceu o seguinte comentário, “a ferramenta poderia ser mais intuitiva”. Porém quando perguntado o que estava faltando, o mesmo não soube responder.

A visualização GTA (Gráfico de Teia e Área), usada para identificar a discrepância entre competências similares, também teve grande aceitação por parte dos participantes. A possibilidade de analisar as competências similares pela diferença de áreas e de iterar sobre os resultados foi o que mais agradou, pois facilitou uma comparação mais profunda entre substitutos e substituído. Como ponto negativo, alguns integrantes destacaram que a visualização de muitas competências ao mesmo tempo, acima de 8, prejudica a percepção das discrepâncias.

As visualizações GLI (Gráfico de Linha para Interesse) e GLC (Gráfico de Linha para Competência), utilizadas para mensurar o grau de interesse num determinado assunto e o grau de conhecimento numa competência, foram bem avaliados pelos participantes. Uma das carências do GCC diz respeito à valoração subjetiva das competências, onde o usuário atribui o valor alto, médio ou baixo às competências que diz possuir. Logo, a possibilidade de visualizar o quanto a competência foi utilizada pelo pesquisador agradou os participantes. A baixa utilização dessas visualizações, conforme ilustrado na Tabela 29, foi decorrente do fato da maior parte dos participantes considerarem as visualizações TM e GTA suficientes na recomendação de substitutos.

A visualização GF (Gráfico de Força), utilizada na exibição de similaridades MBTI, obteve a menor média dentre todas as visualizações avaliadas. Apesar dos participantes aprovarem a utilização da proximidade na percepção de perfis similares, as duas notas baixas (ID1 e ID7) foram decorrentes da refutação ao critério similaridade MBTI. A não-

utilização desta visualização, por alguns participantes, também está relacionada ao fato de considerarem tal critério irrelevante na busca por um substituto.

Comparação entre a busca da FBEE e a do ambiente GCC

A Tabela 30 apresenta os resultados referentes à satisfação do usuário quanto à busca através da FBEE perante a do GCC. A primeira coluna apresenta as perguntas feitas aos participantes. As cinco colunas seguintes mostram as respostas dos participantes, onde MM = Muito melhor; M = Melhor; E = Equiparável; P = Pior e MP = Muito Pior.

Tabela 30 – Resultados referentes à comparação da busca realizada pela FBEE com a realizada através do GCC.

Considerando a busca através da FBEE, compare-a em relação a do GCC quanto:					
	MM	M	E	P	MP
1. A identificação dos critérios relevantes	5	6	1	0	0
2. A localização dos substitutos	7	5	0	0	0
3. A satisfação com os resultados obtidos	8	4	0	0	0
4. Facilidade de uso	9	3	0	0	0

Os entrevistados ficaram satisfeitos com a apresentação dos critérios relevantes utilizados em cada uma das visualizações. Isso foi decorrente da possibilidade de analisar todos os critérios de uma só vez e na mesma tela. Fato que não ocorria no GCC, onde os participantes se viam obrigados a mudarem constantemente de tela, quando não abriam duas telas ao mesmo tempo.

A identificação de possíveis substitutos também foi bem avaliada, a forma rápida e objetiva com que localização dos candidatos passou a ser realizada agradou os participantes. No GCC, era necessário repetir todo o processo de análise para cada possível substituto, tornando o processo de identificação lento e cansativo.

A grande satisfação dos participantes com os resultados obtidos é outro fator importante, coisa que não ocorreu quando os mesmos utilizaram as funcionalidades existentes no GCC. A possibilidade de analisar todos os possíveis candidatos, comparativamente, de uma só vez gerou resultados mais precisos e coerentes dentro de cada grupo, conforme mostrado na Tabela 28. Os participantes também ressaltaram que o

uso de visualizações facilitou o processo de escolha. Vale lembrar que anteriormente a comparação era realizada através de relatórios.

Outro atrativo da FBEE em relação às funcionalidades providas pelo GCC refere-se à facilidade de uso. Os participantes chegaram com maior facilidade ao resultado desejado quando navegaram através de um conjunto finto de visualizações do que quando utilizaram consultas e análise de relatórios.

Nesta seção fizemos à comparação entre as duas abordagens de busca possíveis, na próxima avaliaremos a satisfação dos entrevistados quanto a FBEE como um todo.

5.4.4. Avaliação da FBEE

A Tabela 31 apresenta os resultados obtidos através do questionário (Q5) na avaliação da satisfação com a FBEE, pelo fato deste questionário ser comum a todos participantes, os resultados dos grupos foram unificados. A primeira coluna apresenta as perguntas feitas aos participantes. As cinco colunas seguintes mostram suas respectivas respostas, onde S = Sim; N = Não; R = Razoável e NA = Não se aplica; A última coluna (A) mostra a análise quantitativa realizada, onde uma questão tem resposta S = Sim, N = Não, ou I = Inconclusiva. Para realizar esta análise foi usado o seguinte critério: para uma questão ser considerada Sim ou Não, era preciso ter 80% das respostas dos participantes, sendo que as respostas ‘Não se aplica’ foram desconsideradas, e as respostas ‘Razoável’ foram divididas na metade e somadas tanto as respostas ‘Sim’ como as respostas ‘Não’. Caso contrário a questão era considerada Inconclusiva.

Tabela 31 – Tabela referente às respostas do questionário(Q5)

Sobre a FBEE					
	S	N	R	NA	A
1. Você ficou satisfeito com a ferramenta?	12				S
2. A utilização da FBEE é simples?	12				S
3. A utilização da FBEE é intuitiva?	10		2		S
4. Você conseguiu realizar seu trabalho?	12				S
5. É preciso memorizar muitos passos para utilizar as funcionalidades da FBEE?		12			N
6. Foi fácil aprender a utilizar a FBEE?	11		1		S

Sobre a FBEE					
7. Você acha que a FBEE pode auxiliar em outras atividades? Exemplifique.	3	1		8	I
8. Ao cometer um erro, você consegue voltar e continuar e realizando a sua tarefa?	10		2		S
9. A FBEE possui todas funcionalidades necessárias para realizar a busca por substitutos?	12				S
10. Os elementos disponíveis na interface estão apresentados de forma clara?	11		1		S
11. A interface da FBEE é amigável?	12				S
12. Você gostou da interface da FBEE?	12				S
13. Você encontrou limitações na interface da FBEE? Se sim, liste as limitações encontradas:		11	1		N
14. Todas as funcionalidades utilizadas funcionaram corretamente? Se não, liste quais:	12				S
15. Você gostou de realizar a busca por substitutos através de visualizações?	12				S
16. O uso se visualizações agilizou o processo de busca por substitutos?	12				S
17. Você encontrou alguma limitação das funcionalidades da FBEE? Se sim, quais são elas:		11		1	N
Sobre a experiência de utilização da FBEE					
18. Em geral, você considerou a experiência satisfatória?	12				S
19. Os objetivos do trabalho foram atingidos?	12				S
20. Você conseguiu realizar suas tarefas com sucesso?	12				S
21. Você utilizaria a FBEE para auxiliar no processo de busca por substitutos?	12				S
Comentários extras e sugestões					
Você tem algum comentário extra a acrescentar?					S

A seguir são feitas algumas análises qualitativas no que diz respeito ao que mais agradou na proposta, ao que mais desagradou e ao levantamento de funcionalidades a serem acrescentadas no futuro.

Quando perguntamos se os participantes achavam que a FBEE poderia auxiliá-los em outras atividades, somente 3(três) responderam que “Sim”.

- O primeiro participante (ID1) sugeriu utilizar a ferramenta na substituição de negociadores em eventos dependentes. Dessa sugestão originou-se o trabalho de RODRIGUES *et al.* (2008), no qual técnicas de Visualização da Informação são utilizadas na substituição de negociadores.
- O segundo (ID2) propôs a utilização da FBEE para verificar o interesse das pessoas no projeto.
- O terceiro (ID5) sugeriu utilizá-la no dimensionamento do efetivo dentro de uma empresa.

Um fato que chamou atenção, é que os 3(três) participantes são os únicos que gerenciam equipes de desenvolvimento.

Quando perguntados sobre a interface da FBEE, obtivemos as seguintes respostas:

- “Visualmente fácil de usar”.
- “Gráficos legíveis”.
- “É gráfica, possibilitando a utilização de uma grande variedade de informações”.
- “Todos os recursos de busca acessíveis através de uma única interface”.
- “Informações exibidas de forma coordenada”.
- “A presença de *labels* informativos facilita o processo de busca”.

De uma maneira geral, todos gostaram da interface da FBEE. Entretanto, alguns aspectos desagradaram alguns participantes:

- “Os ícones de informação estavam fora de foco” (ID2 e ID6). Estes ícones são referentes aos sumários que cada ferramenta apresenta para seu entendimento.
- “A coloração do Mapa de Árvores é destoante do ambiente GCC, sugiro uma variação de azul” (ID7).

O problema dos ícones de informação ocorreu com os participantes que utilizaram uma resolução de vídeo inferior a 1280x768, esse problema já foi resolvido. Quanto a

coloração do Mapa de Árvores, decidimos mantê-la, visto que somente um dos participantes reclamou e por assumirmos que a variação de verde mais perceptível que a variação de azul.

Quando perguntados sobre as funcionalidades da FBEE, todos participantes aprovaram o funcionamento e utilidade das mesmas. Entretanto, nenhum comentário adicional foi apresentado.

Os comentários extras e sugestões são sumarizados a seguir:

- A diversidade de opções é um grande atrativo.
- Por ser uma ferramenta visual, a atividade de análise foi facilitada, o que torna a tomada de decisões para um substituto mais rápida.
- Apesar de utilizar “applet”, gráficos e árvore hiperbólica, o sistema tem um bom desempenho.
- A possibilidade de se utilizar mais de um critério ao mesmo tempo agiliza a busca.
- Acho a busca pelo perfil MBTI arriscada, por ser uma análise mais simples do que acontece na vida real (ID1).
- Dificuldade em identificar a rede social dos participantes.

A próxima seção apresenta a conclusão deste estudo.

5.4.5. Conclusão Geral do Estudo Investigativo

O uso das visualizações possibilitou que aos participantes a realização de análises comparativas, intuitivas e simultâneas entre contextos, possíveis substitutos e substituído, reduzindo assim o tempo gasto na busca e facilitando o processo de substituição. Vale lembrar, que muitos participantes utilizaram papel e caneta ou mais de uma tela ao mesmo tempo durante a busca realizada através do GCC.

Alguns participantes refutaram a utilização do perfil MBTI, por considerá-lo irrelevante ou por considerar sua comparação através de um programa pretenciosa demais. Futuramente, pretendemos estudar o perfil MBTI mais profundamente e reavaliar sua relevância na busca por substitutos aptos.

A identificação da rede social dos pesquisadores é feita pela visualização Mapa de Árvores, através dos projetos e comunidades que os pesquisadores participam juntamente. Concordamos que esta identificação deixou um pouco a desejar, entretanto a dissertação de mestrado do aluno MONCLAR (2008) trata deste assunto, podendo futuramente ser adaptada como uma visualização auxiliar a FBEE.

Apesar destas duas últimas observações, a FBEE teve grande aceitação por parte dos participantes, os quais são totalmente favoráveis à implantação da mesma no ambiente GCC.

6. Conclusões e Trabalhos Futuros

Para se encontrar um pesquisador substituto, um grande volume de informação deve ser analisada. Essa sobrecarga faz com que os avaliadores gastem muito tempo processando grandes volumes de dados ou baseiem sua escolha na recomendação de terceiros. Conseqüentemente, na maioria das vezes, o pesquisador escolhido não é o mais indicado. Logo, torna-se necessário o uso de alguma estratégia para organizar e representar essas informações, de modo que os problemas gerados pela sobrecarga de informação sejam minimizados.

Objetivando facilitar o processo de substituição acima descrito, este trabalho propõe a BEE, uma abordagem de busca baseada na substituição através de exemplo. Para isso, a BEE identifica os locais que sofrerão maior impacto após a saída do pesquisador e recomenda candidatos substitutos conforme a similaridade entre suas competências e as do substituído, adicionalmente informações referentes ao perfil MBTI e a capacidade de aprendizado podem ser consideradas. Todos os passos dessa abordagem são apoiados por técnicas de Visualização de Informação. O uso dessas técnicas durante a análise de dados possibilita a agregação da informação e conseqüentemente amplia a cognição, pois permite ao ser humano reparar em detalhes e detectar padrões, trazendo à tona informações que seriam difíceis de analisar e entender.

A BEE foi implementada no GCC para validarmos nossa abordagem. No contexto do GCC, esse trabalho substitui a busca textual por pesquisadores, na qual os resultados eram retornados na forma de relatórios, o que impossibilitava a análise de toda informação relevante à identificação de substitutos potenciais.

Nosso estudo de casos mostrou que a utilização dessa abordagem trouxe grandes benefícios à busca por substitutos no GCC. O uso das visualizações possibilitou aos avaliadores a realização de análises comparativas, intuitivas e simultâneas das competências do contexto de atuação do pesquisador, dos possíveis substitutos e da pessoa a ser substituída, reduzindo assim o tempo gasto na busca e facilitando o processo de substituição. Outra característica propiciada foi o filtro por contexto, ou seja, a utilização do contexto do pesquisador a ser substituído para reduzir o volume de informação a ser analisada pelo usuário. Neste caso, somente as competências dos pesquisadores, comuns ao

contexto da busca e de seus respectivos itens, são utilizadas na comparação de similaridades.

Entretanto, em um contexto mais abrangente (SREs), a principal contribuição desse trabalho está na utilização de diferentes tipos de visualizações para cada faceta do perfil do pesquisador, onde: o Mapa de Árvores exibe o contexto de atuação e a importância do pesquisador; os Gráficos de Teia e Área exibem as competências do pesquisador relevantes ao contexto de busca; o Gráfico de Força ilustra as características do perfil MBTI e por fim os Gráficos de Linha exibem a evolução e conhecimento do pesquisador num determinado interesse e competência respectivamente. Essa característica facilita consideravelmente o processo de substituição e não foi encontrada em nenhum dos SREs analisados.

A partir de uma análise crítica sobre a abordagem proposta e sua implementação, puderam ser identificadas diversas limitações e uma série de trabalhos futuros, apresentados a seguir:

- Uso da solução proposta em uma situação de cenário real. Pretendemos avaliar a usabilidade desta abordagem através de um estudo de caso em um cenário científico real, utilizando grupos de domínios diferentes de nossa universidade, para depois expandir este estudo em diferentes universidades brasileiras. Lembramos que nem todas as informações utilizadas em nosso estudo de caso eram reais. Ao aplicarmos a abordagem a cenários reais, podem surgir outras demandas de visualizações e funcionalidades que não tenham sido previstas na implementação da abordagem nem indicadas durante o estudo de casos realizado.
- Utilização da BEE em ambientes não-científicos. Da forma que essa abordagem foi proposta, esperamos que apenas modificações pontuais – ou seja: o conjunto de informações analisadas – sejam necessárias para integrarmos a abordagem proposta a outros tipos de ambiente. Desta maneira, poderemos analisar as reais diferenças na busca de um substituto em um ambiente acadêmico e em um ambiente comercial.
- Suporte a análise da reputação dos candidatos substitutos. CRUZ *et al.* (2007) propõe um Modelo de Sistema para Reputação em Comunidades Virtuais, que procura coletar informações sobre as participações

(mensuradas de forma quantitativa) e contribuições (mensuradas de forma qualitativa) dos usuários em cada área de interesse em que eles atuam, para utilizar como indicativos das suas reputações. Acreditamos que a inclusão deste modelo trará grande melhoria aos resultados da BEE, pois não só a quantidade, mas também a qualidade, das contribuições providas pelo candidato substituto passarão a ser consideradas pelo avaliador durante o processo decisório.

- **Melhoria da Análise da Rede Social.** Atualmente a identificação da rede social dos pesquisadores é feita pela visualização Mapa de Árvores, através dos projetos e comunidades que os pesquisadores participam juntamente. Entretanto, pelo fato de muitas vezes ser imperceptível, essa identificação é desconsiderada pela pessoa responsável por encontrar o pesquisador substituto. MONCLAR (2008) propõe a identificação e balanceamento de redes sociais de forma a garantir o fluxo contínuo de conhecimento no ambiente científico. Essas redes são exibidas através de um *gráfico radial*, no qual os nós representam pesquisadores e as arestas o relacionamento entre eles. Logo, a adição dessa visualização à BEE facilitará a identificação dos relacionamentos de cada pesquisador, suprimindo assim a limitação hoje existente em nossa abordagem.
- **Concordância do pesquisador em ser recomendado.** Afinal, para que a recomendação seja bem sucedida, o pesquisador deve estar disposto a migrar de área. Uma idéia inicial é que a BEE entre em contato de maneira automática com os substitutos candidatos por email. No protótipo implementado no GCC, essa comunicação ocorre através de mensagens internas ao sistema, fóruns e convites.
- **Verificação de Disponibilidade.** A disponibilidade do pesquisador é um fator relevante durante a escolha. Normalmente, pesquisadores relacionados a um menor número de projetos e comunidades costumam ser mais receptivos a participarem de novas comunidades e projetos. Logo, uma maneira de checar essa disponibilidade traria grandes benefícios a nossa abordagem.

- Enriquecimento do perfil dos candidatos. Pretendemos realizar estudos relacionados à identificação de novos itens a serem inseridos no perfil dos candidatos substitutos que aumentariam a precisão dos resultados retornados.

Como trabalho futuro, tem-se ainda a busca por novas estruturas visuais. Apesar dos usuários terem aprovado as visualizações utilizadas, sabemos que sempre um novo tipo de visualização pode surgir e, conseqüentemente, se adequar melhor a apresentação dos resultados. Por exemplo, caso encontrássemos uma visualização que nos possibilitasse analisar todas as facetas do perfil em uma mesma tela e com a mesma eficácia de apresentação ela seria utilizada, pois reduziria o número de passos para se alcançar o resultado desejado.

Referências Bibliográficas

- ADAM, P., MARC, S., 2006, “Contrasting Portraits of Email Practices: Visual approaches to reflection and analysis”. In: *Proceedings of the Working Conference on Advanced Visual Interfaces*, pp. 389 – 395, Venezia, Italy, May.
- AHLBERG, C., SHNEIDERMAN, B., 1994, “Visual Information Seeking: Tight Coupling of Dynamic Query Filters with Starfield Displays”. In: *Proceedings of the CHI'94, ACM Conference on Human Factors in Computing Systems*, pp. 313-317, New York, USA, April.
- AHLBERG, C., WISTRAND, E., 1995, “IVEE: An Information Visualization and Exploration Environment”. In: *Proceedings of the Symposium on Information Visualization*, pp. 66-73, Atlanta, Georgia, USA, October.
- AUTOFOCOS PERSONAL, 2007, “AUTOFOCOS PERSONAL”. Disponível em <http://www.aduna.biz>. Acessado em 03/12/2007.
- BARBOSA, C. E., OLIVEIRA, J., MAIA, L., et al., 2007, “Using Recommendation Systems for Explicit Knowledge Dissemination and Profiling Identification for Scientific and Engineering Contexts”. In: *11th International Conference on Computer Supported Cooperative Work in Design*, pp. 715-721, Melbourne, Australian, April.
- BARROS, M. O., WERNER, C. M. L., TRAVASSOS, G. H., 2002, "Um Estudo Experimental sobre a Utilização de Modelagem e Simulação no Apoio à Gerência de Projetos de Software". In: *XIX Simpósio Brasileiro de Engenharia de Software (SBES)*, Uberlândia, MG, Brasil, Outubro.
- BECKS, A., REICHLING, T., WULF, V., 2004, “Expertise Finding: Approaches to Foster Social Capital”. In: *Social Capital and Information Technology*, Huysman, M. and Wulf, V. (eds.). MIT Press, Cambridge, pp. 333-354.
- BENJAMIN, B. B., SHNEIDERMAN, B., WATTENBERG, M., 2002, “Ordered and quantum treemaps: Making effective use of 2D space to display hierarchies”, *ACM Transactions on Graphics (TOG)*, v.21, n. 4 (October), pp. 833-854.
- BERNERS, L. T., HENDLER, J., LASSILA, O., 2001, “The Semantic Web”, In: *Scientific American*, v. 284, n. 5(August), pp. 35-43. <http://www.sciam.com/article.cfm?id=the-semantic-web>
- BIRK, A., 1997, “Modeling the Application Domains of Software Engineering Technologies”. In: *Proceedings of the 12th IEEE International Conference on Automated Software Engineering*, pp. 291-292, Incline Village, NV, USA, November.
- BOMFIM, E. P., 2005, *THOTH: Uma Ferramenta para Reutilização de Processos Científicos*. M.Sc. , COPPE/UFRJ, Rio de Janeiro, Brasil.

- BONHARD, P., SASSE, A., HARRIES, C., 2007, “The Devil You Know Knows Best How Online Recommendations Can Benefit From Social Networking”. In: *Proceedings of the 21st British HCI Group Annual Conference*, pp.77-88, Lancaster, UK, September.
- BRAUTIGAM, M., 1996, *Applying Information Visualization Techniques to Web Navigation*, Dissertation Proposal. University of California, Santa Cruz, CA.
- BRULS, D. M., HUIZING, C., VAN WIJK, J. J., 2000, “Squarified Treemaps”. In: *Proceedings of the joint Eurographics and IEEE TCVG Symposium on Visualization*, pp. 33-42, Amsterdam, Netherlands, May.
- CAMPANHOLA, C., SILVA, J. G., 2000, “Desenvolvimento Local e a Democratização dos Espaços Rurais”. In: *Cadernos de Ciência e Tecnologia*, Brasília, v. 17, n. 1 (Março), pp. 11-40.
- CAPES, 2008, Disponível em: <http://www.capes.gov.br>. Acessado em 06/06/2008.
- CARD, S. K., MACKINLAY, J. D., SHNEIDERMAN, B., 1999. *Readings in Information Visualization: Using Vision to Think*. 1st ed, San Francisco, USA, Morgan Kaufman Publishers Inc.
- CAVA, R. A., LUZZARDI, P. R. G., FREITAS, C. M. D. S., 2002, “The Bifocal Tree: a Technique for the Visualization of Hierarchical Information Structures”. In: *15th Workshop on Human Factors in Computer Systems (IHC)*, Fortaleza, Brazil, October.
- CELMAR, S. G., 2006, *Exploração de bases de dados de ambientes de Educação a Distância por meio de Ferramentas de Consulta Apoiadas por Visualização de Informação*. D.Sc., UNICAMP, São Paulo, Brasil.
- CHEN, C., 1999, *Information Visualization and Virtual Environments*. 1st ed, London, UK, Springer Verlag.
- CHEN, C., 2005, “Top 10 Unsolved Information Visualization Problems”, *IEEE Computer Graphics and Applications*, v. 25, n. 4 (July), pp. 12-16.
- CHEN, C., BÖRNER, K., 2002, “Visual Interfaces to Digital Libraries: Motivation, Utilization, and Socio-technical Challenges”. In: *Visual Interfaces to Digital Libraries*, v. 2539, *Lecture Notes in Computer Science*. Springer-Verlag, pp.1-9.
- CNPQ, 2007, “CNPQ“. Disponível em: <http://www.cnpq.br/areasconhecimento/index.htm>. Acessado em: 01/12/2007.

- CRAPO, W. A., LAURIE, B. W., WILLIAM, A. W., et al., 2000, "Visualization and The Process of Modeling: A Cognitive-theoretic View". In: *Proceedings of the 6th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining KDD*, pp. 218-226, Boston, USA, August.
- CRASWELL, N., HAWKING, D., VERCOUSTRE, A. M., et al., 2001, "P@NOPTIC Expert: Searching for Experts not just for Documents". Poster in AusWeb'01, Australian, April. <http://ausweb.scu.edu.au/aw01/papers/edited/vercoustre/paper.html>,
- CRUZ, C. C. P., MOTTA, C. L. R., SANTORO, F. M., 2007, "RECOP: Um Modelo para Reputação em Comunidades de Prática". In: *XVIII Simpósio Brasileiro de Informática na Educação SBIE 2007*, pp. 200-211, São Paulo, Brasil, Novembro.
- EICK, S. G., COX K. C., TAOSONG, H. E., 1996, "3D Geographic Network Displays". *ACM Sigmod Record*, v. 25, n. 4 (December). pp. 50-54.
- E-QF, 2008. "FRAUNHOFER E-QUALIFICATION FRAMEWORK". Disponível em: <http://www.e-qf.de>. Acessado em: 09/07/2008.
- FERREIRA, C. B. R., NASCIMENTO, H. A. D., 2005, "Visualização de Informações – Uma Abordagem Prática". In: *Anais do XXV Congresso da Sociedade Brasileira de Computação*, pp. 1262-1312, São Leopoldo, RS, Brasil, Julho.
- FLUIT, C., HARMELEN, V. F., SABOU, M., 2005, "Ontology-based Information Visualization: Towards Semantic Web Applications", In: Geroimenko, V. (ed), *Visualising the Semantic Web*, 2nd ed., chapter 3, New York, USA, Springer Verlag.
- FRY, B., 2007, "REVISIONIST", Disponível em: <http://acg.media.mit.edu/people/fry/revisionist>. Acessado em: 03/12/2007
- GOOGLE, 2007, "GOOGLE". Disponível em: <http://www.google.com.br>. Acessado em: 26/11/2007.
- GOSLING, J., JOY, B., STEELE, G., et al., 2005, *The Java Language Specification*. 3ª ed., Prentice Hall PTR.
- GROKKER, 2007, "GROKKER". Disponível em: <http://www.grokker.com>. Acessado em: 20/11/2007.
- HAREL, D., 1987, Statecharts: "A Visual Formalism for Complex Systems". *Science of Computer Programming*, v. 8, n. 3 (June), pp.231-274.
- HAVRE, S., HETZLER, B., NOWELL, L., 1999, "ThemeRiver: In Search of Trends, Patterns, and Relationships". In: *Proceedings of the IEEE Symposium on Information Visualization*, pp.115-123, San Francisco, USA, October.

- HEER, J., CARD, S. K., LANDAY, A., 2005, "PREFUSE: A Toolkit for Interactive Information Visualization". In: *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 421-430, Portland, Oregon, USA, April.
- JFREECHART, 2008, "JFREECHART". Disponível em: <http://www.jfree.org/jfreechart/>. Acessado em: 14/03/2008.
- JUDELMAN, G. B., 2004, *Knowledge Visualization. Problems and Principles for Mapping the Knowledge Space*. M.Sc., University of Lübeck, Germany.
- JUNG, C. G., 1971, *Psychological Types*. 6th ed, London, Princeton University Press.
- KAWAMURA, V., 2006, *Inteligência Competitiva para Instituições de Ensino e Pesquisa*, M.Sc., COPPE/UFRJ, Rio de Janeiro, Brasil.
- KARTOO, 2007, "KARTOO". Disponível em: <http://www.kartoo.net>. Acessado em: 12/12/2007.
- KEIM, D. A., HAO, M. C., DAYAL, U., et al., 2002, "Pixel Bar Charts: a Visualization Technique for very large Multi-Attribute Data Sets". *Information Visualization*, v. 1, n. 1(March), pp.20-34.
- KEIRSEY, D., BATES, M., 1984, *Please Understand Me – Character & Temperament Types*. 5th ed, Prometheus Nemesis Book Company.
- KOHONEN, T., 1989, *Self-Organization and Associative Memory*, 3rd ed, Springer-Verlag, Berlin.
- KREJCI, D., OLIVEIRA, J., SOUZA, J. M., 2008a, "The Use of Visualization and Mining Techniques for Analysis and Recommendation of Communities in the Brazilian Scientific Scenario", *IADIS Web Based Communities 2008 (WBC 2008) Conference*, 21, Amsterdam, Netherlands, 22-27 July.
- KREJCI, D., OLIVEIRA, J., SOUZA, J. M., 2008b, "The Use of Visualization and Mining Techniques for Research Recommendation in the Brazilian Scientific Scenario", *IADIS Visual Communication 2008: Creative Industries, Photography and Culture (VC 2008) Conference.*, 13, Amsterdam, Netherlands, 26 July.
- KRISHNAN, M., BOHN, S. J., COWLEY, W., et al., 2007, "Scalable Visual Analytics of Massive Textual Datasets." In: *Proceedings of the IEEE International Parallel and Distributed Processing Symposium (IPDPS 2007)*, pp. 1-10, Long Beach, CA, USA, March.
- KROEGER, O., THUESEN, J. M., 1988, *Type Talk: The 16 Personality Types that Determine How We Live, Love, and Work*, 1st ed, Dell Publishing, New York.

- LAMPING, J., RAO, R., 1995, "The Hyperbolic Browser: A Focus + Context Technique for Visualizing Large Hierarchies". In: *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 401-408. San Francisco, USA, May.
- LENGLER, R., EPPLER, M., J., 2007, "Towards A Periodic Table of Visualization Methods for Management". In: *LASTED Proceedings of the Conference on Graphics and Visualization in Engineering (GVE 2007)*, pp. 72-78, Clearwater, Florida, USA, January.
- LONG, J. G., AND DOROTHY E. D., 1994, "Ultra-Structures: A Design Theory for Complex Systems and Processes". *Communications of the ACM*, v. 38, n. 1(January), pp.103-120.
- MACKINLAY, J. D., 1986, "Automating the Design of Graphical Presentations of Relational Information". *ACM Transactions on Graphics (TOG)*, v. 5, n. 2(April), pp. 110-141.
- MATTOX, D., MAYBURY, M., MOREY, D., 1999, "Enterprise Expert and Knowledge Discovery". In: *Proceedings of the HCI International '99 (the 8th International Conference on Human-Computer Interaction) on Human-Computer Interaction: Communication, Cooperation, and Application Design*, pp.303-307, Mahwah, NJ, USA, August.
- MAYER, R. E., 2001, *Multimedia Learning*. 2nd ed, Cambridge University Press, New York.
- MYERS, I. B., 1980. *Gifts Differing: Understanding Personality Type*. 1st ed, Davies-Black Publishing.
- MILLER, G., 1956, "The Magical Number Seven, Plus or Minus Two: Some Limits on our Capacity for Processing Information". *The Psychological Review*, v. 63, n. 2(March.), pp. 81-97.
- MITRE, 2008. "MITRE". Disponível em: www.mitre.org. Acessado em 05/01/2008.
- MONCLAR, R. S., 2008, *Análise e Balanceamento de Redes Sociais Científicas*, M.Sc., COPPE/UFRJ, Rio de Janeiro, RJ, Brasil.
- NORMAN, D. A., 1993, *Things that make us smart: Defining human attributes in the age of the machine*. 1st ed, Addison-Wesley Publishing Company, New York.
- NORTH, C., 2005, "Information Visualization", In: G. Salvendy (ed). *Handbook of Human Factors and Ergonomics*, 3rd ed, chapter 46, John Wiley & Sons, New York.
- OLIVEIRA, K.M., 1999, *Modelo para Construção de Ambientes de Desenvolvimento de Software Orientados a Domínio*, D.Sc., COPPE/UFRJ, Rio de Janeiro, RJ, Brasil.

- OLIVEIRA, J., SOUZA, J. M., MIRANDA, R., et al., 2005, "GCC: An Environment for Knowledge Management in Scientific Research and Higher Education Centers". *In: I-Know - 5th International Conference on Knowledge Management, Special Track on Knowledge Sharing in Research and Higher Education*, pp. 145-156, Graz, Austria, June.
- OLIVEIRA, J., SOUZA, J. M., MIRANDA, et al., 2006, "GCC: A Knowledge Management Environment for Research Centers and Universities". *In: Frontiers of WWW Research and Development - APWeb 2006*, v. 3841, *Lecture Notes in Computer Science*. Hei-delberg, Germany: Springer-Verlag, pp. 652-667.
- OLIVEIRA, J., 2007, *METHEXIS: Uma Abordagem de Apoio à Gestão do Conhecimento para Ambientes de "E-Science"*, D.Sc., COPPE/UFRJ, Rio de Janeiro, RJ, Brasil.
- OPTE PROJECT, 2007, "OPTE PROJECT". Disponível em: <http://www.opte.org/maps/>. Acessado em 23/11/2007.
- PNL, 2007, "PNL" Disponível em: <http://www.pnl.gov/InfoViz/technologies.html>. Acessado em 12/12/2007.
- PREECE, J., 2000, *Online Communities: Designing Usability, Supporting Sociability*. 1st ed. New York, John Wiley & Sons, Inc.
- PREFUSE, 2008, "PREFUSE". Disponível em: <http://www.prefuse.org>. Acessado em 12/04/2008.
- REVISIONIST, 2007, "REVISIONIST". Disponível em: <http://acg.media.mit.edu/people/fry/revisionist>. Acessado em 13/12/2007.
- ROCHA, H. V., BARANAUSKAS, M. C. C., 2003, *Design e Avaliação de Interfaces Humano-Computador*, 1^a ed. Campinas, SP: Nied/Unicamp.
- RODRIGUES, S., SAMPAIO, J. O., SOUZA, J. M., 2004, "Competence Mining for Virtual Scientific Community Creation". *International Journal of Web Based Communities*, v.1, n.1 (July), pp. 90-102.
- RODRIGUES, S., KREJCI, D., SOUZA, J. M., 2008, "Negotiation Supported Through Visualization Methods", *In: XXXIV Conferencia Latinoamericana de Informática (CLEI 2008)*, 45166, Santa Fé, Argentina. 8-12 de Setembro.
- SANTOS, G., VILLELA, K., ROCHA, A. R., 2003, "SAPIENS: uma ferramenta para descrição e recuperação de competências em uma organização", *I Workshop de TI e Gerência do Conhecimento*, Fortaleza, CE.
- SCAIFE, M., ROGERS, Y., 1996, "External Cognition: How Do Graphical Representations Work?", *International Journal of Human-Computer Studies*, v. 45, n. 2 (August), pp. 185-213.

- SHNEIDERMAN, B., 1992, "Tree Visualization with TreeMaps: 2D Space-Filling Approach". *ACM Transactions on Graphics*, v. 11, n. 1(January), pp. 92-99.
- SHNEIDERMAN, B., 1994, "Dynamic Queries for Visual Information Seeking". *IEEE Software*, v.11, n. 6(November), pp. 70-77.
- SHNEIDERMAN, B., 1996, "The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations". In: *Proceedings of the IEEE Conference on Visual Languages*, pp. 336-343, Columbus, Ohio, USA , September.
- SHNEIDERMAN, B., BENJAMIN B. B., 2003, *The Craft of Information Visualization: Readings and Reflections*, 2nd ed, Morgan Kaufmann Publishers Inc. San Francisco, CA, USA.
- SHNEIDERMAN, B., PLAISANT, C., 2005, *Designing the User Interface: Strategies for Effective Human-Computer Interaction*, 4th ed, Addison-Wesley.
- SINNZEUG, 2007, "SINNZEUG". Disponível em: <http://www.sinnzeug.de>. Acessado em: 03/12/2007.
- SOLINGEN, V. R., BERGHOUT, E., 1999, *The Goal / Question / Metric Method: A Practical Guide for Quality Improvement of Software Development*, McGraw-Hill.
- SPENCE, R., 2001, *Information Visualization*, 1st ed, Addison Wesley Publisher.
- STOLTE, C., TANG, D., HANRAHAN, P., 2002, "Polaris: A System for Query, Analysis and Visualization of Multi-dimensional Relational Databases". *IEEE Transactions on Visualization and Computer Graphics*, v. 8, n. 1(March), pp. 52-65.
- TERVEEN, L. G., MCDONALD, D. W., 2005, "Social Matching: A Framework and Research Agenda". *ACM Transactions on Computer-Human Interaction (ToCHI)*, v. 12, n. 3(September), pp. 401-434.
- TUFTE, E. R., 2001. *Envisioning Information*. 2nd ed, Graphics Press.
- VILLELA, K., TRAVASSOS, G.H., ROCHA, A.R., 2001, "Ambientes de Desenvolvimento de Software Orientados à Organização", In: *Proceedings of IDEAS'2001*, San Jose, Costa Rica, April.
- VIVACQUE, A. S., LIEBERMAN, H., 2000, "Agents to Assist in Finding Help". In: *Proceedings of the Conference on Computer Human Interaction (CHI 2000)*, pp. 65-72, Hague, Netherlands, April.

- VIVACQUA, A. S., FERREIRA, M. S., SOUZA D.,K., et al., 2006, “Navigator: Uma Ferramenta para Colaboração Oportunística”. Relatório Técnico ES-704/06. Rio de Janeiro: Universidade Federal do Rio de Janeiro.
- VIVISIMO, 2007, “VIVISIMO”. Disponível em: www.vivisimo.com. Acessado em: 12/12/2007.
- WARE, C., 2004, *Information Visualization: Perception for Design*. 2nd ed, Morgan Kaufmann Publisher, San Francisco.
- WATTENBERG, M., 1999, “Visualizing the Stock Market”. In: *Proceedings of the Conference on Human Factors in Computing Systems (CHI 99)*, pp.188-189, Pittsburgh, Pennsylvania, May.
- WATTENBERG, M., 2002, “Arc Diagrams: Visualizing Structure in Strings”. In: *Proceedings of the IEEE Symposium on Information Visualization (InfoVis'02)*, pp.110-116. Washington, DC, USA, October.
- WEBSOM, 2007, “WEBSOM”. Disponível em: <http://websom.hut.fi/websom/comp.ai.neural-nets-new/html/root.html>. Acessado em 12/12/2007.
- WEHREND, S., LEWIS, C., 1990, “A Problem-Oriented Classification of Visualization Techniques”. In: *Proceedings of the IEEE Visualization (Vis'90)*, pp. 139-143, San Francisco, California, October.
- WICKENS, C. D., HOLLANDS, J. G., 2000, *Engineering Psychology and Human Performance*, 3rd ed, Prentice-Hall.
- WOHLIN, C., RUNESON, P., HÖST, M., OHLSSON, M.C., REGNELL, B., WESSLÉN, A., 2000, *Experimentation in Software Engineering: an Introduction*, Kluwer Academic Publishers, Norwell, MA
- YANG, J., WARD, M., RUNDENSTEINER E., 2003. “Interactive Hierarchical Displays: A General Framework for Visualization and Exploration of Large Multivariate Data Sets”. *Computers and Graphics Journal*, v. 27, n. 2(April), pp.265-283.
- YAO, C., PENG, B., HE, J, et al., 2005, “CNDS Expert Finding System for TREC2005”. In: *Proceedings of the 14th Text Retrieval Conference*, pp.303-307, Gaithersburg, Maryland, November.
- YIMAM-SEID, D., KOBASA, A., 2003, “Expert Finding Systems for Organizations: Problem and Domain Analysis and the DEMOIR Approach”. *Journal of Organizational Computing & Electronic Commerce*, v.13, n.1 (November). Pp.1-24

YUPENG, F., XIANG, R., ZHANG, M., et al., 2006, "A PDD-Based Searching Approach for Expert Finding in Intranet Information Management". *In: Information Retrieval Technology*, v. 4182, *Lecture Notes in Computer Science*, Heidelberg, Germany: Springer-Verlag , pp.43-53.

ZHOU, M., FEINER, S., 1998, "Visual Task Characterization for Automated Visual Discourse Synthesis". *In: Proceedings of the ACM Human Factors in Computing Systems Conference (CHI'98)*, pp 392-399. Addison-Wesley Publishing, New York, USA. April

Anexo I – Questionário de Avaliação do Estudo

Questionário 1

Questões de Caracterização dos Participantes

*Assinale uma das opções disponíveis para cada questão.

Dados Pessoais

1. Nome:
2. E-mail:
3. Ano de Nascimento:
4. Sexo: () masculino () feminino

Formação Acadêmica

5. Titulação:
 - () Graduação Incompleta
 - () Graduação Completa
 - () M.Sc. Incompleto
 - () M.Sc. Completo
 - () D.Sc. ou Ph.D. Incompleto
 - () D.Sc. ou Ph.D. Completo
6. Área de Formação:
7. Instituição onde recebeu o título mais recente:
8. Você gerencia ou lidera algum projeto? () Sim () Não
9. Se sim, a sua equipe é composta de quantas pessoas?
 - () 0 - 10
 - () 11 - 20
 - () 21 - 30
 - () 31 - 40
 - () 41 - 50
 - () +50

Experiência Profissional

10. Área de Atuação:
11. Instituição em que exerce função:
12. Principais atividades:

Questionário 2 – GCC X FBEE (Módulo de Projetos)

***Somente os participantes que receberam o treinamento no módulo de projetos devem responder o questionário 2.**

Sobre a organização e a estrutura do trabalho

1. Você ficou satisfeito com a apresentação do trabalho?
() sim () não () razoável () não se aplica
2. Você ficou satisfeito com o treinamento referente ao ambiente GCC?
() sim () não () razoável () não se aplica
3. Você ficou satisfeito com o treinamento referente a FBEE?
() sim () não () razoável () não se aplica
4. Você ficou satisfeito com o treinamento como um todo?
() sim () não () razoável () não se aplica

Busca por substitutos através do GCC

Após a saída do “Participante 1” do “Projeto 1”, informe:

5. Que critérios foram considerados na busca por um substituto:

6. Que funcionalidades do ambiente foram utilizadas na busca por um substituto:

7. Quanto tempo foi gasto na busca por um substituto:
Hora de Início:
Hora de Término:
Total:

8. O(s) possíveis substitutos encontrados, em ordem de relevância, e as razões para sua escolha:

9. O seu grau de satisfação quanto aos substitutos encontrados, e as razões:

Busca por substitutos através da FBEE

10. Os critérios abaixo foram considerados na busca por um substituto em um projeto. Assinale os que você considera válidos e atribua-lhes uma pontuação de 1-5:
() iteratividade no projeto (tarefas, materiais, publicações, contribuições);
() ser gerente do projeto;

- menor discrepância entre os valores das competências requeridas pelo projeto, as do participante a ser substituído e do substituto;
- grau de conhecimento na atividade que era realizada pelo usuário a ser substituído;
- interesse na atividade que era realizada pelo usuário a ser substituído;
- comparação com o perfil psicológico do substituído com os possíveis substitutos (MBTI).

11. Qual(is) possíveis substitutos foram encontrados, ordene por sua relevância. Porquê essa ordem foi escolhida?

12. O tempo gasto na busca por um substituto?

Hora de Início:

Hora de Término:

Total:

13. Avalie as visualizações utilizadas e justifique?(1-5)

- TreeMap(determinação da importância nos projetos)

- Gráfico de Teia e Área (discrepância de conhecimento)

- Gráfico de Linha (nível de conhecimento numa competência)

- Gráfico de Linha (aprendizado na competência de interesse)

- Gráfico de Força (similaridade de perfil MBTI)

Considerando a busca por substitutos na FBEE, compare-a com a do ambiente GCC quanto a(o):

14. Localização dos itens relevantes:

muito melhor melhor equiparável pior muito pior

Justifique:

15. Localização do substituto:

muito melhor melhor equiparável pior muito pior
Justifique:

16. Satisfação com os resultados obtidos:
 muito maior maior equiparável menor muito menor
Justifique:

17. Facilidade de uso:
 muito maior maior equiparável menor muito menor
Justifique:

18. Tempo gasto na localização do substituto:
 muito maior maior equiparável menor muito menor
Justifique:

Questionário 3 – GCC X FBEE (Módulo de Comunidades)

***Somente os participantes que receberam o treinamento no módulo de comunidades devem responder o questionário 3.**

Sobre a organização e a estrutura do trabalho

1. Você ficou satisfeito com a apresentação do trabalho?
() sim () não () razoável () não se aplica
2. Você ficou satisfeito com o treinamento referente ao ambiente GCC?
() sim () não () razoável () não se aplica
3. Você ficou satisfeito com o treinamento referente a FBEE?
() sim () não () razoável () não se aplica
4. Você ficou satisfeito com o treinamento como um todo?
() sim () não () razoável () não se aplica

Busca por substitutos através do GCC

Após a saída do “Participante 1” da “Comunidade 1” informe:

5. Que critérios foram considerados na busca por um substituto:

6. Que funcionalidades do ambiente foram utilizadas na busca por um substituto:

7. Quanto tempo foi gasto na busca por um substituto:

Hora de Início:

Hora de Término:

Total:

8. O(s) possíveis substitutos encontrados, em ordem de relevância, e as razões para sua escolha:

9. O seu grau de satisfação quanto aos substitutos encontrados, e as razões:

Busca por substitutos através da FBEE

10. Os itens abaixo foram considerados na busca por um substituto em uma comunidade, assinale os que você considera válido e atribua uma pontuação de 1-5 aos mesmos:

() iteratividade na comunidade (materiais, publicações, eventos, enquetes, fórun, reunião eletrônica, links);

- ser moderador da comunidade;
- ser ponte entre outras comunidades;
- menor discrepância entre os valores das competências requeridas pela comunidade, as do participante a ser substituído e do substituto;
- grau de conhecimento no assunto do usuário a ser substituído;
- interesse nos assuntos discutidos pelo usuário a ser substituído;
- comparação entre o perfil psicológico (MBTI)

11. Qual(is) possíveis substitutos foram encontrados, ordene por sua relevância. Porquê essa ordem foi escolhida?

12. O tempo gasto na busca por um substituto?

Hora de Início:

Hora de Término:

Total:

13. Avalie as visualizações utilizadas e justifique?(1-5)

- TreeMap(determinação da importância nos projetos)

- Gráfico de Teia e Área (discrepância de conhecimento)

- Gráfico de Linha (nível de conhecimento numa competência)

- Gráfico de Linha (aprendizado na competência de interesse)

- Gráfico de Força (similaridade de perfil MBTI)

Considerando a busca por substitutos na FBEE, compare-a com a do ambiente GCC quanto a(o):

14. Localização dos itens relevantes:

muito melhor melhor equiparável pior muito pior

Justifique:

15. Localização do substituto:

muito melhor melhor equiparável pior muito pior

Justifique:

16. Satisfação com os resultados obtidos:

muito maior maior equiparável menor muito menor

Justifique:

17. Facilidade de uso:

muito maior maior equiparável menor muito menor

Justifique:

18. Tempo gasto na localização do substituto:

muito maior maior equiparável menor muito menor

Justifique:

Questionário 4 – GCC X FBEE (Contexto Organizacional)

***Somente os participantes que receberam o treinamento no contexto organizacional devem responder o questionário 4.**

Sobre a organização e a estrutura do trabalho

1. Você ficou satisfeito com a apresentação do trabalho?
() sim () não () razoável () não se aplica
2. Você ficou satisfeito com o treinamento referente ao ambiente GCC?
() sim () não () razoável () não se aplica
3. Você ficou satisfeito com o treinamento referente a FBEE?
() sim () não () razoável () não se aplica
4. Você ficou satisfeito com o treinamento como um todo?
() sim () não () razoável () não se aplica

Busca por substitutos através do GCC

Após a saída do “Participante 1” da “Instituição” informe:

5. Que critérios foram considerados na busca por um substituto:

6. Que funcionalidades do ambiente foram utilizadas na busca por um substituto:

7. Quanto tempo foi gasto na busca por um substituto:
Hora de Início:
Hora de Término:
Total:

8. O(s) possíveis substitutos encontrados, em ordem de relevância, e as razões para sua escolha:

9. O seu grau de satisfação quanto aos substitutos encontrados, e as razões:

Busca por substitutos através da FBEE

10. Os itens abaixo foram considerados na busca por um substituto na organização, assinale os que você considera válido e atribua uma pontuação de 1-5 aos mesmos:
() participação em comunidades

- iteratividade em comunidades (materiais, publicações, eventos, enquetes, fórun, reunião eletrônica, links);
- ser moderador de comunidades;
- ser ponte entre comunidades;
- participação em projetos;
- iteratividade em projetos (tarefas, materiais, publicações, contribuições);
- ser gerente de projetos;
- menor discrepância entre os valores das competências do participante a ser substituído e do substituto;
- grau de conhecimento no assunto do usuário a ser substituído;
- grau de interesse nos assuntos discutidos pelo usuário a ser substituído;
- comparação entre o perfil psicológico (MBTI)

11. Qual(is) possíveis substitutos foram encontrados, ordene por sua relevância. Porquê essa ordem foi escolhida?

12. O tempo gasto na busca por um substituto?

Hora de Início:

Hora de Término:

Total:

13. Avalie as visualizações utilizadas e justifique?(1-5)

- TreeMap (determinação da importância nos projetos)

- Gráfico de Teia e Área (discrepância de conhecimento)

- Gráfico de Linha (nível de conhecimento numa competência)

- Gráfico de Linha (aprendizado na competência de interesse)

- Gráfico de Força (similaridade de perfil MBTI)

Considerando a busca por substitutos na FBEE, compare-a com a do ambiente GCC quanto a(o):

14. Localização dos itens relevantes:

muito melhor melhor equiparável pior muito pior

Justifique:

15. Localização do substituto:

muito melhor melhor equiparável pior muito pior

Justifique:

16. Satisfação com os resultados obtidos:

muito maior maior equiparável menor muito menor

Justifique:

17. Facilidade de uso:

muito maior maior equiparável menor muito menor

Justifique:

18. Tempo gasto na localização do substituto:

muito maior maior equiparável menor muito menor

Justifique:

Questionário 5 – Avaliação da FBEE

Sobre a usabilidade da FBEE

1. Você ficou satisfeito com a ferramenta?
() sim () não () razoável () não se aplica

2. A utilização da FBEE é simples?
() sim () não () razoável () não se aplica

3. A utilização da FBEE é intuitiva?
() sim () não () razoável () não se aplica

4. Você conseguiu realizar o trabalho utilizando a FBEE?
() sim () não () razoável () não se aplica

5. É preciso memorizar muitos passos para utilizar as funcionalidades da FBEE?
() sim () não () razoável () não se aplica

6. Foi fácil aprender a utilizar a FBEE?
() sim () não () razoável () não se aplica

7. Você acha que a FBEE pode te auxiliar em outras atividades?
() sim () não () razoável () não se aplica
Exemplifique:

8. Ao cometer um erro, você consegue voltar e continuar realizando a sua tarefa?
() sim () não () razoável () não se aplica

9. A FBEE possui todas as funcionalidades necessárias para a realização da busca por substitutos?
() sim () não () razoável () não se aplica
Se não, cite alguma:

10. Os elementos disponíveis na interface estão apresentados de forma clara?
() sim () não () razoável () não se aplica

11. A interface da FBEE é amigável?
() sim () não () razoável () não se aplica

12. Você gostou da interface da FBEE?

sim não razoável não se aplica

Liste aspectos os positivos e negativos:

13. Você encontrou alguma limitação na interface da FBEE?

sim não razoável não se aplica

Se sim, liste as limitações identificadas:

14. Todas as funcionalidades utilizadas funcionaram corretamente?

sim não razoável não se aplica

Se não, liste as funcionalidades:

15. Você gostou de realizar a busca por substitutos através de visualizações?

sim não razoável não se aplica

16. O uso de visualizações agilizou o processo de busca por substitutos?

sim não razoável não se aplica

17. Você encontrou alguma limitação nas funcionalidades da FBEE?

sim não razoável não se aplica

Se sim, liste as limitações identificadas:

Sobre a Experiência de utilização da FBEE

18. Em geral, você considerou a experiência satisfatória?

sim não razoável não se aplica

19. Os objetivos do trabalho foram atingidos?

sim não razoável não se aplica

20. Você conseguiu realizar suas tarefas com sucesso?

sim não razoável não se aplica

21. Você utilizaria a FBEE para auxiliar no processo de busca por substitutos?

sim não razoável não se aplica

Comentários Extras e Sugestões

22) Você tem algum comentário extra a acrescentar?

Anexo II – Modelo de Dados Utilizado pela FBEE

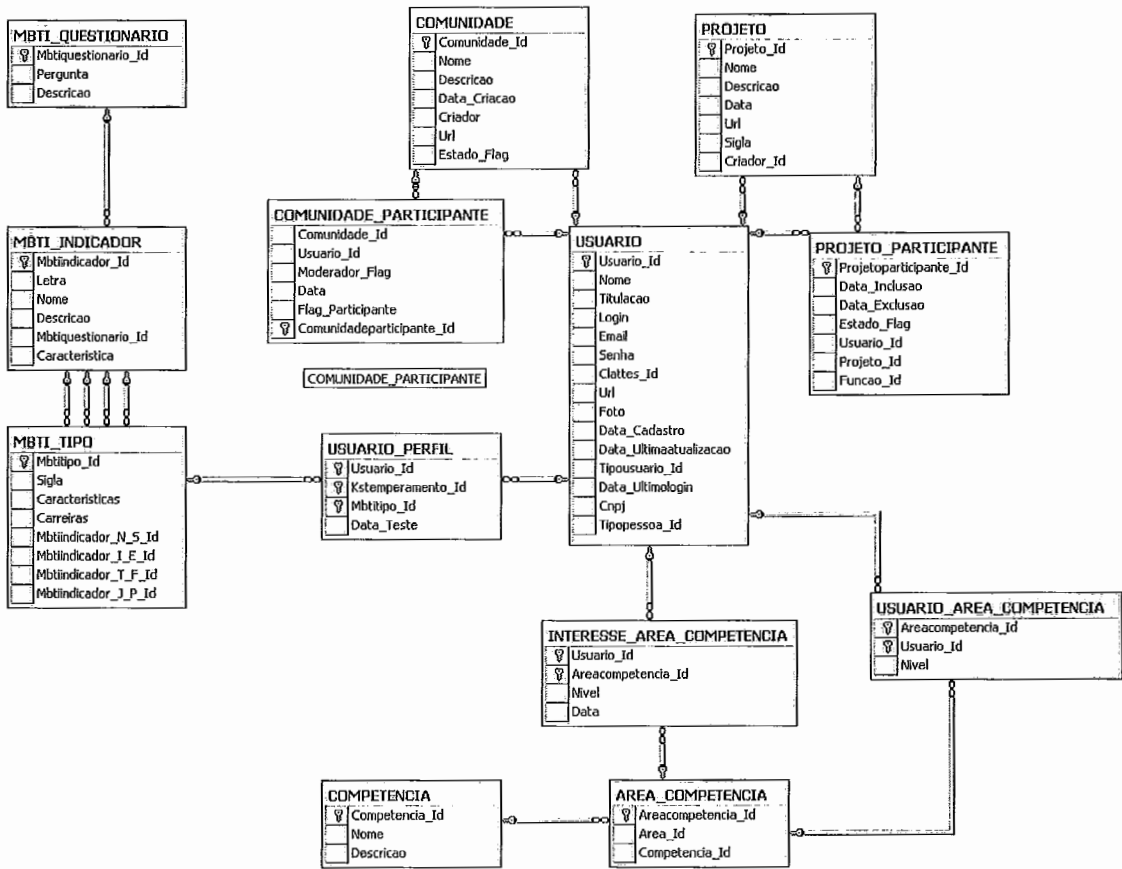


Figura 41 – Informações do usuário no contexto organizacional.

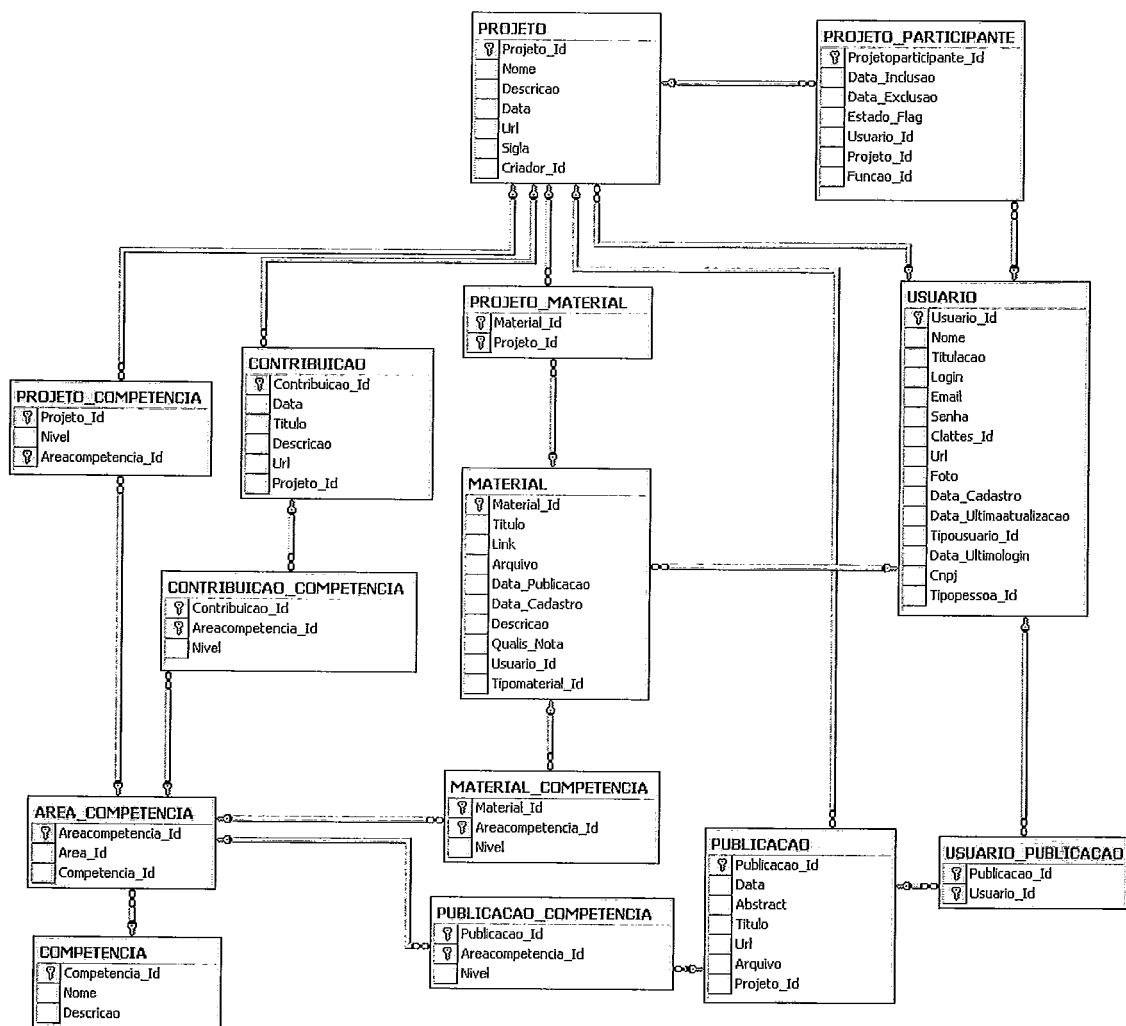


Figura 42 – Informações do usuário no contexto projeto.

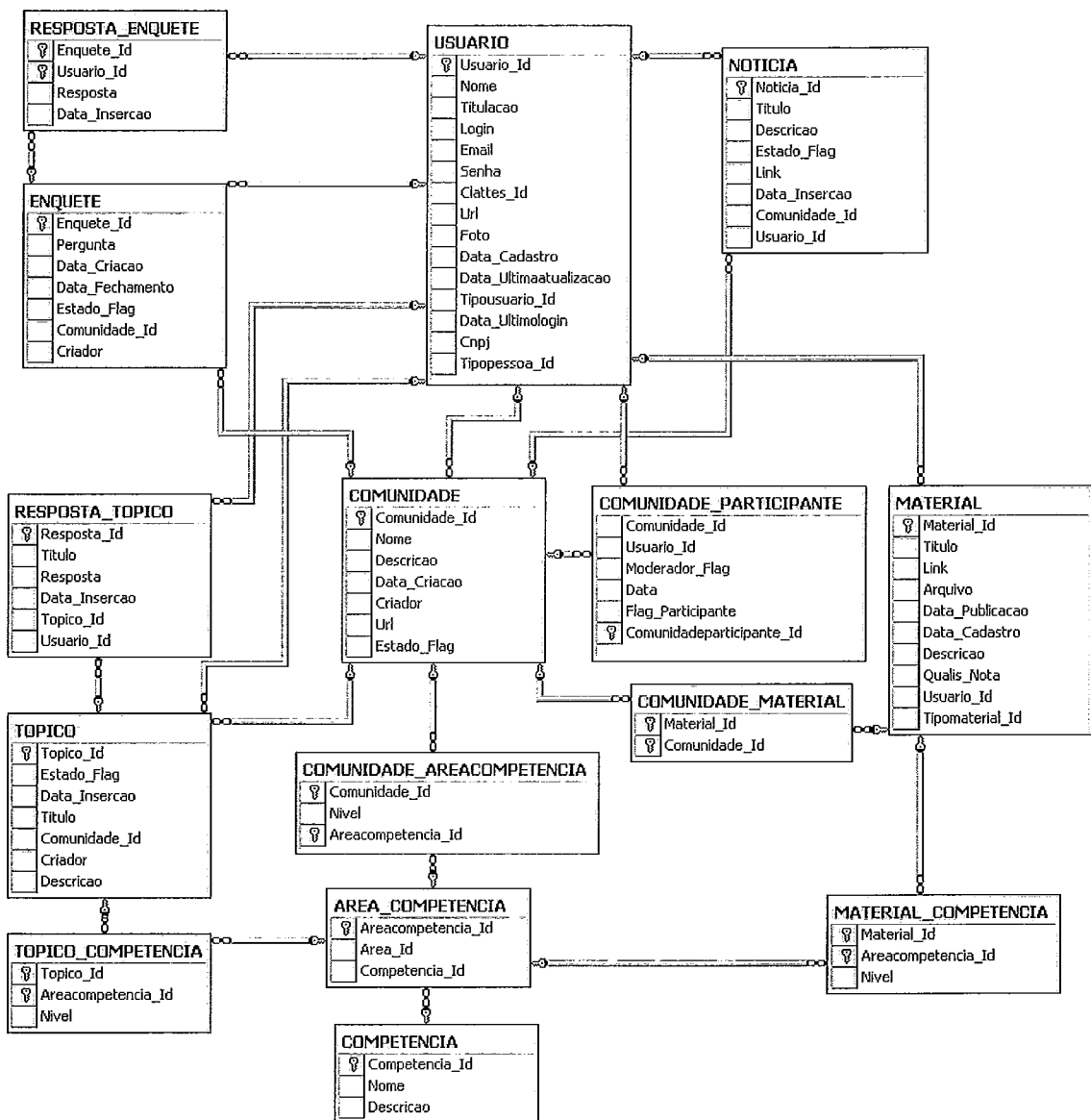


Figura 43 – Informações do usuário no contexto comunidade