



**UNIVERSIDADE SALVADOR – UNIFACS
PROGRAMA DE PÓS-GRADUAÇÃO EM SISTEMAS E
COMPUTAÇÃO
MESTRADO ACADÊMICO EM SISTEMAS E COMPUTAÇÃO**

MARCOS LAPA DOS SANTOS

**FINDYOURHELP: UM MÓDULO PARA APOIAR A BUSCA POR
ESPECIALISTAS NO AMBIENTE VIRTUAL DE APRENDIZAGEM
MOODLE**

Salvador
2010

MARCOS LAPA DOS SANTOS

**FINDYOURHELP: UM MÓDULO PARA APOIAR A BUSCA POR
ESPECIALISTAS NO AMBIENTE VIRTUAL DE APRENDIZAGEM
MOODLE**

Dissertação apresentada à Universidade Salvador, como parte das exigências do Curso de Mestrado Acadêmico em Sistemas e Computação, para obtenção do título de Mestre.

Orientadora: Prof^a Laís do Nascimento Salvador

Co-Orientadora: Prof^a Daniela Soares Cruzes

Salvador
2010

FICHA CATALOGRÁFICA

Elaborada pelo Sistema de Bibliotecas da Universidade Salvador – UNIFACS

Santos, Marcos Lapa dos

Findyourhelp: um módulo para apoiar a busca por especialistas no ambiente virtual de aprendizagem moodle/ Marcos Lapa dos Santos. - 2010.

135 f. : il.

Dissertação (Mestrado) - Universidade Salvador – UNIFACS. Mestrado em Sistemas de Computação, 2010.

Orientador: Prof. Dr. Laís do Nascimento Salvador

Co-Orientadora: Prof^a Daniela Soares Cruzes

1. Ambientes virtuais. 2. Tecnologia educacional. 3. Educação a Distancia. I. Salvador, Laís do Nascimento, orient. II. Cruzes, Daniela Soares, orient. III. Título.

CDD: 005.1

FINDYOURHELP: UM MÓDULO PARA APOIAR A BUSCA POR ESPECIALISTAS NO AMBIENTE VIRTUAL DE APRENDIZAGEM MOODLE

Dissertação aprovada como requisito parcial para obtenção do grau de Mestre em Sistemas e Computação, Universidade Salvador – UNIFACS, pela seguinte banca examinadora:

Laís Nascimento do Salvador – Orientadora _____
Doutora em Engenharia Elétrica, Universidade de São Paulo - USP
Universidade Federal da Bahia – UFBA

José Maria Nazar David _____
Doutor em Engenharia de Sistemas e Computação, COPPE/Universidade Federal do Rio de Janeiro - UFRJ
Universidade Salvador - UNIFACS

Vaninha Vieira dos Santos _____
Doutora em Ciência da Computação, Universidade Federal de Pernambuco - UFPE
Universidade Federal da Bahia - UFBA

Salvador

maio de 2010.

AGRADECIMENTOS

A vida nos põe em contato com pessoas que, surpreendentemente, se tornam tão importantes, que de repente não podemos imaginar como seria nosso percurso se não as tivéssemos encontrado pelo caminho. Desta forma, aproveito o momento para agradecer àqueles que foram fundamentais para que eu me inspirasse e completasse mais este ciclo de vida. Agradeço, portanto:

Ao criador deste universo que habitamos, pois sem Ele nada seria possível. Obrigado por mais essa oportunidade de evolução moral e intelectual.

À minha mãe MARY, por me incentivar de todas as formas possíveis a seguir em frente diante de todos os obstáculos que a vida me ofereceu. E também por ser batalhadora, guerreira e não desistir de mim nunca.

Aos meus torcedores mais fiéis: TIA NORMA, TIO GIL, TIA ZEZA, TIA ELIANA e MARQUINHO, TIA CEIÇA, TIO GERALDO e família, TIA EDNA e família, minha madrinha querida TIA BELA, minha avó VITÓRIA, vovó ADÉLIA, TIA JÚLIA, TIO JORGE e família, MIRIAN, LÍCIA, LÉIA e família e meus primos-irmãos TICO, LOLY e RÓ com suas respectivas famílias. Todos me acompanham desde que nasci e me trazem sempre a certeza de que tenho vários cantos neste mundo (pelo menos nesta Bahia) para voltar caso um dia seja necessário.

Às minhas meninas ÍRIS e GIULIA, que são o meu constante combustível desde que chegaram em minha vida. Penso que sem vocês nada disso hoje faria sentido. E à minha grande “pequena” JANAINA, namorada, esposa, mãe de minhas pequenas e, acima de tudo, a grande responsável por eu começar e terminar este trabalho, visto que tem a Educação correndo em suas veias e desde o primeiro momento me fez acreditar que isto seria possível, estando ao meu lado e me mostrando o que é ser “uma companheira para todas as horas”.

Aos meus amigos e demais integrantes da família que de certa forma suportaram minhas ausências e mesmo assim continuaram me incentivando sempre que a conversa enveredava mais uma vez para o assunto: Mestrado.

Aos que partiram antes de ver este trabalho se concretizar, mas tenho certeza de que estariam orgulhosos de mim: TIO MARCO, VÔ SALVADOR, VÓ NILZETE, VÓ EDELVIRA e TIA GRACINHA.

Aos professores, meus amigos, que participaram do estudo de caso respondendo aos questionários aplicados e que me ajudaram de alguma forma a analisar e melhorar a ferramenta aqui produzida.

Por fim, um agradecimento especial às minhas orientadoras: professora LAÍS e professora DANIELA, pela paciência de todos os dias, correções dos erros, sugestões, modificações necessárias e reuniões via Skype tão enriquecedoras para o meu processo produtivo. E aos professores componentes desta banca, VANINHA e JOSÉ MARIA, por aceitarem avaliar e ajudar a ajustar o trabalho neste último e não menos importante momento de intervenção.

RESUMO

O presente trabalho analisa a adoção de técnicas de mineração de textos aliada a conceitos de gestão do conhecimento como forma de apoiar a busca por especialistas em fóruns de discussão no ambiente Moodle. A solução, intitulada FindYourHelp, foi desenvolvida como um módulo adicional ao Moodle e permite a execução de um algoritmo de categorização de textos no momento da postagem de uma mensagem. Como forma de validar a aplicabilidade da ferramenta, foi desenvolvido um estudo de viabilidade onde se buscou a execução da solução de forma a analisar as mensagens de três disciplinas distintas pertencentes a duas Instituições de Ensino Superior da Bahia. Os resultados do estudo revelaram que a solução proposta obteve êxito na maioria dos casos analisados. Os três docentes responsáveis pelas disciplinas, participaram da avaliação dos resultados apontados pelo FindYourHelp e confirmaram a sua confiabilidade no processo de identificar os especialistas automaticamente em fóruns de discussão acadêmicos.

Palavras-chave: Gestão do Conhecimento. Socialização do Conhecimento. Mineração de Textos. Busca de Especialistas. Ambientes Virtuais de Aprendizagem. Moodle. FindYourHelp.

ABSTRACT

This study examines the adoption of text mining techniques with the concepts of knowledge management to support the search for experts in the discussion forums on Moodle environment. The solution, called FindYourHelp, was developed as an additional Moodle plugin and allows the execution of a text categorization algorithm at the time of a message posting. In order to validate the tool applicability, a feasibility study was developed where the main intent was to execute the solution in order to analyze the messages of three separate disciplines from two institutions of higher education in Bahia. The study results showed that the proposed solution was successful in most analyzed cases. The three disciplines professors, participated in the results evaluation indicated by FindYourHelp and confirmed its reliability in the process of automatically identifying experts in academic discussion forums.

Palavras-chave: Knowledge Management. Knowledge Socialization. Text Mining. Expert Find. Learning Management Systems. Moodle. FindYourHelp.

LISTA DE ILUSTRAÇÕES

Figura 1 - Espiral do Conhecimento.....	17
Figura 2 - Buscadores de especialistas integrados a sistemas de informações organizacionais.....	27
Figura 3 - Sistemas de Busca por Especialistas.....	29
Figura 4 - Modelo de serviços do AulaNet classificados em função do modelo 3C.....	32
Figura 5 - Parte da arquitetura do ambiente AulaNet.....	33
Figura 6 - Exemplo de tela de acompanhamento do Dokeos.....	40
Quadro 1 - Comparativo de ambientes virtuais de aprendizagem.....	41
Figura 7 - Passos que compõem o processo de KDD.....	46
Figura 8 - Etapas da mineração de textos.....	49
Figura 10 - Associando documentos a categorias.....	51
Quadro 2 - Representação de um Vetor de Termos de um Documento (MAGALHÃES, 2008).....	54
Quadros 3 e 4 - Exemplo de representação de um texto em relação aos seus termos.....	55
Figura 11 - Modelo do Espaço Vetorial.....	55
Figura 12 - Categorização de mensagens sem a utilização de um método de ponderação.....	57
Figura 13 - Arquitetura do FindYourHelp integrada ao AVA Moodle.....	67
Figura 14 - Diagrama de Atividades: Inserir novo termo/categoria.....	69
Figura 15 - Diagrama de Atividades: Análise de uma postagem no fórum.....	71
Figura 16 - Estrutura de tabelas de sinônimos do FindYourHelp.....	73
Figura 17 - Diagrama de atividades: Identificação e exibição de especialistas.....	75
Figura 18 - Tela de exibição da árvore de categorias do módulo FindYourHelp.....	77
Figura 19 - Tela de inserção/edição de categorias/tópicos do módulo FindYourHelp.....	78
Figura 20 - Tela de exclusão de categoria/tópico do módulo FindYourHelp.....	79
Figura 21 - Lista de especialistas identificados pelo módulo FindYourHelp.....	79
Figura 22 - Tela de validação de especialistas.....	81
Quadro 5 - Disciplinas analisadas pelo estudo de viabilidade da ferramenta.....	84
Figura 23 - Processamento de mensagens importadas do FindYourHelp.....	85
Figura 24 - Hierarquia de termos da disciplina de Linguagem de Programação Orientada a Objetos I.....	91
Figura 25 - Mensagens Categorizadas da Disciplina 1.....	94
Figura 26 - Mensagens Descartadas da Disciplina 2.....	97
Figura 27 - Mensagens Categorizadas da Disciplina 2.....	99
Figura 28 - Mensagens descartadas da Disciplina 3.....	102
Quadro 6 - Comparativo entre os resultados das disciplinas analisadas.....	104

SUMÁRIO

1 INTRODUÇÃO	12
1.1 METODOLOGIA ADOTADA	13
1.2 ESTRUTURA DO TRABALHO	14
2 GESTÃO DO CONHECIMENTO E A BUSCA POR ESPECIALISTAS	15
2.1 TIPOS DE CONHECIMENTO	15
2.2 A ESPIRAL DO CONHECIMENTO.....	17
2.3 O PROCESSO DE SOCIALIZAÇÃO DO CONHECIMENTO	19
2.4 TIPOS DE FERRAMENTAS PARA AUXÍLIO À SOCIALIZAÇÃO DO CONHECIMENTO.....	20
2.4.1 Portais de Conhecimento	21
2.4.2 Sistemas de Trabalho em Grupo (Groupware)	22
2.4.3 Sistemas de Interações Síncronas	23
2.4.4 Ambientes para aprendizagem on-line	23
2.5 BUSCA POR ESPECIALISTAS	25
2.6 CONSIDERAÇÕES FINAIS	30
3 AMBIENTES VIRTUAIS DE APRENDIZAGEM	31
3.1 AULANET	31
3.2 MOODLE	34
3.3 BLACKBOARD.....	35
3.4 TELEDUC.....	37
3.5 DOKEOS.....	39
3.6 COMPARATIVO ENTRE OS AMBIENTES PESQUISADOS	40
3.7 UM OLHAR PROPOSITIVO SOBRE AMBIENTES VIRTUAIS DE APRENDIZAGEM	42
4 MINERAÇÃO DE TEXTOS	45
4.1 MINERAÇÃO DE DADOS E KDD.....	46
4.2 MINERANDO DADOS EM FORMA DE TEXTO LIVRE	47
4.3 ANÁLISE SEMÂNTICA (ENTENDENDO UM TEXTO).....	50
4.4 CATEGORIZANDO TEXTOS.....	51
4.5 CONSIDERAÇÕES FINAIS	59
5 FINDYOURHELP: BUSCANDO ESPECIALISTAS DENTRO DE UM AVA	60
5.1 REQUISITOS FUNCIONAIS	61
5.2 CARACTERÍSTICAS TÉCNICAS	65
5.3 ARQUITETURA DO MÓDULO FINDYOURHELP	66
5.4 PRINCIPAIS FUNCIONALIDADES.....	68
5.4.1 Exibir e atualizar a hierarquia de categorias	68
5.4.2 Analisar as postagens dos fóruns	70
5.4.3 Identificar especialistas para uma categoria escolhida	74
5.5 INTERAÇÃO COM O USUÁRIO DO AVA MOODLE	76
5.5.1 Configurando a árvore de categorias	76
5.5.2 Listando os especialistas	79
5.5.3 Validando os especialistas identificados	80
5.6 CONSIDERAÇÕES FINAIS	82
6 FINDYOURHELP: ESTUDO DE VIABILIDADE DA FERRAMENTA	83
6.1 O AMBIENTE DE APLICAÇÃO DO ESTUDO	83
6.2 A COLETA DOS DADOS	84
6.3 ANÁLISE: CRITÉRIOS E VALIDAÇÕES	86
6.4 LIMITAÇÕES DO ESTUDO E AMEAÇAS ÀS VALIDAÇÕES	88
6.5 MONTAGEM DAS HIERARQUIAS DE CATEGORIAS	90
6.6 ANÁLISE DOS RESULTADOS	91
6.6.1 Disciplina: Linguagem de Programação Orientada a Objetos I	92
6.6.2 Disciplina: Projeto Avançado de Sistemas II	95
1.1.1. Disciplina: Tecnologias Interativas Aplicadas à Educação	100
6.7 CONSIDERAÇÕES FINAIS.....	104
7 CONCLUSÃO	106
7.1 CONTRIBUIÇÕES	107
7.2 TRABALHOS FUTUROS	108
7.3 LIMITAÇÕES E CONSIDERAÇÕES FINAIS.....	110
REFERÊNCIAS	111
APÊNDICE A – Documentos utilizados no estudo de viabilidade	115

APÊNDICE B – Código para importação do OpenthesauruS	117
APÊNDICE C – Sinais Gráficos Removidos	122
ANEXO A – Respostas aos questionários do estudo.....	123

1 INTRODUÇÃO

Os avanços promovidos pela adoção em grande escala das Tecnologias da Informação e Comunicação (TIC) nos diversos setores da sociedade possibilitaram a geração e armazenamento de dados sobre serviços, produtos, pessoas e demais fontes de informação.

O conhecimento passou a ser um elemento-chave para o sucesso das instituições, tornando necessário que estas dispensem uma atenção especial ao tratamento oferecido aos dados e informações que trafegam todos os dias em seus domínios, sobretudo aos seres humanos envolvidos neste fluxo (NONAKA; TAKEUCHI, 1997).

Dentro deste cenário, percebe-se a adoção dos ambientes virtuais de aprendizagem (AVA) pelas Instituições de Ensino Superior (IES), como ferramentas importantes na elaboração e re-elaboração do conhecimento em rede.

Em meio ao grande movimento de informações nestes ambientes, encontram-se aquelas pessoas que se destacam dentro de grupos de discussão ou pesquisa, por conhecerem profundamente determinados assuntos de interesse coletivo. Estas são geralmente percebidas como especialistas dentre as demais e por isso, tornam-se as mais requisitadas dentro da equipe onde contribuem.

Docentes que conseguem mobilizar a participação dos seus estudantes em rede, através de recursos tecnológicos como fóruns ou chats, percebem, em sua maioria, uma grande mudança na postura do estudante, agora visto como co-ator na “criação da comunicação e da aprendizagem.” (SILVA, 2003, p. 263).

Um problema no cenário atual de ferramentas AVA é que a maior parte das informações veiculadas é em formato de texto. Faz-se necessário então uma investigação de técnicas que podem apoiar docentes e discentes no processo de aprendizagem e reconhecimento de especialistas.

Uma abordagem explorada neste trabalho é o uso de técnicas de Mineração de Textos em mensagens contidas em um AVA, aliado à aplicação de conceitos de socialização do conhecimento.

Delineia-se então o principal objetivo deste trabalho, que é analisar a adoção de técnicas de mineração de texto, como apoio ao processo de socialização do

conhecimento produzido em grupos de discussão acadêmicos, que utilizam como ambiente de aprendizagem virtual a ferramenta Moodle.

Os objetivos específicos que norteiam a execução deste trabalho são:

- a) Atualização do estado da arte sobre Gestão do Conhecimento, AVA, Busca de Especialistas e Mineração de Textos;
- b) Discussão sobre técnicas para categorização de textos de curto comprimento, tais como mensagens de fóruns de discussão;
- c) Proposta de um módulo, intitulado **FindYourHelp**, que permita a identificação automática e busca de especialistas no AVA Moodle;
- d) Condução de um estudo de viabilidade com a finalidade de analisar a aplicabilidade do **FindYourHelp** em grupos de discussão relacionados a três disciplinas ministradas em Instituições de Ensino Superior.

1.1 METODOLOGIA ADOTADA

Nesta pesquisa foi construído um módulo adicional ao AVA Moodle, com a finalidade de analisar e validar a aplicabilidade da solução proposta. A ferramenta permite a interceptação de mensagens postadas no recurso de fórum do ambiente de forma a categorizá-las e pontuar seus autores em um *ranking* de participantes mais ativos (considerados neste trabalho como especialistas).

Visando o cumprimento do objetivo do presente trabalho, foi realizado um estudo bibliográfico sobre os conceitos de gestão do conhecimento, ferramentas de busca por especialistas e mineração de textos, com um enfoque nas técnicas de categorização automática de textos, bem como um estudo comparativo sobre diversos AVA com notória adoção por Instituições de Ensino Superior no Brasil.

Como forma de complementação a este estudo, foi aplicado um questionário com professores de disciplinas que utilizam recursos de fóruns virtuais em sua prática pedagógica, de forma a levantar informações que fossem relevantes para a confecção do módulo **FindYourHelp**.

Pode-se destacar ainda a execução de um estudo de caso, o qual se utilizou das mensagens postadas em fóruns de discussão de três disciplinas distintas como base para execução e validação da solução proposta. Duas destas disciplinas são integrantes do curso de graduação em Sistemas de Informação do Centro Universitário da Bahia (ESTACIO/FIB) e a terceira faz parte do curso de pós-

graduação em Metodologia e Didática do Ensino Superior da Universidade Católica de Salvador (UCSAL). Os três professores responsáveis pelas disciplinas atuaram como usuários do módulo com a finalidade de analisá-lo à luz de sua vivência profissional.

Desta forma foi aplicado um questionário semi-estruturado com estes docentes, a fim de obter um retorno contendo as impressões gerais sobre o uso e resultados apontados pela solução proposta.

1.2 ESTRUTURA DO TRABALHO

Os três capítulos que seguem a introdução contêm a fundamentação teórica e revisão da literatura relacionada ao tema. Os dois capítulos subseqüentes discutem o trabalho desenvolvido nesta dissertação. Por fim tem-se a conclusão, as referências eletrônicas e bibliográficas utilizadas, em seguida os anexos e apêndices referentes aos questionários aplicados e ao material

Portanto, o presente trabalho está organizado da seguinte forma:

O capítulo 1 descreve a motivação, os objetivos, metodologia adotada pela pesquisa, o seu escopo e a organização do presente documento.

O capítulo 2 discute a teoria de Gestão do Conhecimento tendo como foco principal o processo de socialização do conhecimento definido por Nonaka e Takeuchi (2007), ao final do capítulo é abordado o conceito de busca por especialistas e suas principais ferramentas de mercado.

O capítulo 3 faz um estudo comparativo entre ambientes virtuais de aprendizagem, tendo em vista levantar informações relevantes para a formulação da proposta de solução aqui desenvolvida.

O capítulo 4 discute o conceito de mineração de textos e suas principais técnicas para categorização automática

O capítulo 5 descreve a ferramenta **FindYourHelp** de forma detalhada e sua implementação.

O capítulo 6 faz um estudo de viabilidade e aplicabilidade da solução proposta através da análise de mensagens contidas nos fóruns de três disciplinas que se utilizaram destes recursos para complementação da prática pedagógica docente.

O capítulo 7 apresenta as conclusões deste trabalho, bem como suas contribuições, perspectivas de trabalhos futuros e limitações.

2 GESTÃO DO CONHECIMENTO E A BUSCA POR ESPECIALISTAS

2.1 TIPOS DE CONHECIMENTO

O tema conhecimento ocupa muitos espaços de discussão na sociedade contemporânea e dada a sua relevância nos diversos cenários sociais, atualmente se revela como objeto de reflexão de pesquisadores de diversas áreas.

O que é conhecimento? Qual é o conhecimento mais eficiente ou que gera maior produtividade para uma empresa? Como transformar informação em conhecimento? Qual é o papel dos acadêmicos frente ao processo de elaboração e socialização do conhecimento em uma sociedade? Como capturar e formalizar o conhecimento de especialistas para uso futuro? Estas e outras perguntas orientam o olhar de muitos pensadores, em especial, gestores e educadores, na direção da reflexão acerca da produção de saberes e seu sentido em um espaço social.

Setzer (1999) define conhecimento como sendo “uma abstração interior, pessoal, de alguma coisa que foi experimentada por alguém”. Já Davenport e Prusak (2003, p. 6) são mais completos quando afirmam que o conhecimento é:

uma mistura fluida de experiência condensada, valores, informação contextual e *insight* experimentado, a qual proporciona uma estrutura para avaliação e incorporação de novas experiências e informações. Ele tem origem e é aplicado na mente dos conhecedores. **Nas organizações, ele costuma estar embutido não só em documentos ou repositórios, mas também em rotinas, processos, práticas e normas organizacionais. (grifo nosso).**

Essas definições permitem que se perceba o conhecimento como sendo mais complexo do que os dados ou as informações, pois este precisa ser avaliado, contextualizado e interpretado a partir de uma experiência prévia do seu conhecedor.

Os autores Nonaka e Takeuchi (1997) trazem em sua obra uma visão interessante de que o conhecimento pode ser dividido em dois tipos principais:

- a) Tácito, ou aquele conhecimento dito pessoal, que envolve as habilidades do indivíduo, específico ao contexto e, portanto, difícil de ser formulado e comunicado.

- b) Explícito, o qual pode ser entendido como o conhecimento “codificado”, transmissível em linguagem formal e sistemática.

Fialho e outros (2006) complementa o pensamento sobre estes dois tipos de conhecimento afirmando que o conhecimento tácito é aquele que se obtém com a prática e trata-se de algo difícil de ser articulado na linguagem formal. Já o conhecimento explícito é aquele que pode ser transmitido formalmente, e é mais facilmente compartilhado entre pessoas.

Davenport e Prusak (2003) chamam atenção para o fato de que existe uma relativa dificuldade de se absorver e transferir o conhecimento dependendo do tipo de conhecimento envolvido. Esses autores afirmam que o conhecimento explícito pode ser incluído em procedimentos e mantido em bancos de dados ou documentos, ganhando assim uma razoável confiabilidade em sua transferência, já o conhecimento tácito depende muito de um intenso contato pessoal, o que pode dificultar a sua transferência.

Tais autores chamam a atenção também para os fatores culturais inibidores da transferência de conhecimento, os quais são chamados de atritos em seu livro. Esses fatores podem retardar ou até mesmo impedir a transferência e os autores propõem formas de superação.

Dentre estes atritos, destacam-se: a falta de confiança mútua, que pode ser superada com a ocorrência de encontros como reuniões face a face para a construção dos relacionamentos e confiança entre as partes; diferenças culturais, vocabulários e quadros de referência, que podem ser superados estabelecendo-se um consenso através de discussão, publicação, educação, trabalho em equipe e rodízio de funções; o status e a recompensa, atribuídos geralmente aos possuidores do conhecimento. Neste caso é necessário pensar em incentivos baseados no compartilhamento e avaliação de desempenho das pessoas; a intolerância em relação a erros ou necessidade de ajuda, onde a superação passa por uma mudança de visão, a qual percebe e recompensa os erros criativos e a colaboração.

Percebe-se, portanto, a natureza humana da ação de gerenciar o conhecimento, onde as pessoas envolvidas possuem seus interesses individuais, que podem interferir direta e negativamente na dinâmica de transferência e difusão do conhecimento.

Pode-se então afirmar, que a gestão do conhecimento é compreendida na condição de área que estuda as formas de captura, compartilhamento, transferência

e transformações do conhecimento e “trata da prática de agregar valor à informação e distribuí-la, tendo como tema central o aproveitamento dos recursos existentes na empresa.” (FIALHO e outros, 2006, p. 84).

2.2 A ESPIRAL DO CONHECIMENTO

Os autores Nonaka e Takeuchi (1997) desenvolveram uma teoria que relaciona os quatro modos de conversão do conhecimento, baseados na teoria da existência dos dois tipos citados anteriormente (tácito e explícito), sendo eles: 1) a conversão do conhecimento tácito em tácito, que recebe o nome de socialização; 2) a conversão do conhecimento tácito em explícito, que recebe o nome de externalização; 3) a conversão do conhecimento explícito em explícito, que recebe o nome de combinação; 4) a conversão do conhecimento explícito em tácito, que recebe o nome de internalização.

As interações promovidas durante o processo de criação do conhecimento dentro das organizações, são chamadas pelos autores supracitados de Espiral do Conhecimento, a qual explica que o conhecimento é convertido em seus diferentes modos devido à indução por diversos fatores, tais como: diálogos, reflexões coletivas, interações, o “aprender fazendo”, dentre outros.

A seguir, pode-se verificar uma figura que representa esta idéia da Espiral do Conhecimento de Nonaka e Takeuchi (1997):

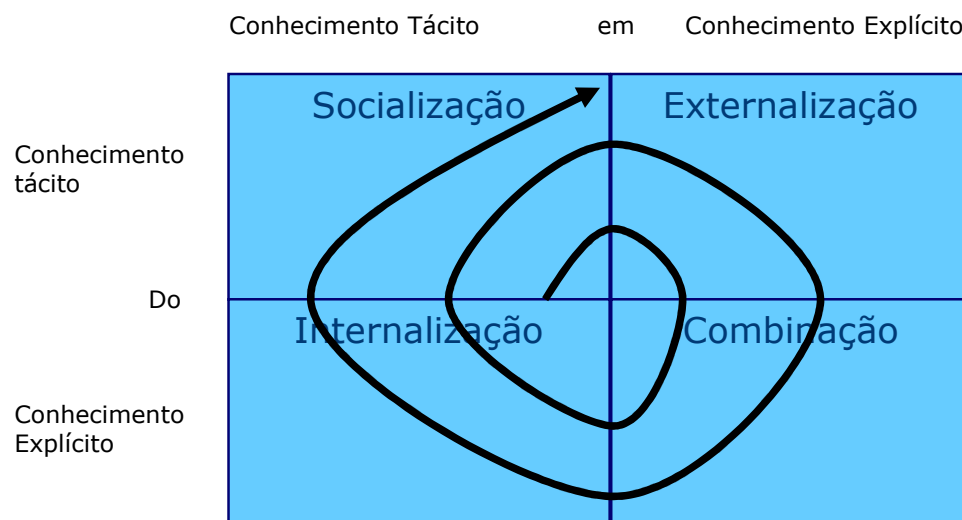


Figura 1 - Espiral do Conhecimento

Nota: Adaptado de Nonaka e Takeuchi (1997, p. 80).

Observando esta espiral é possível compreender que ela se refere a um processo contínuo e retro-alimentado, onde cada conversão pode ser vista como um passo que inicia uma próxima conversão.

Os autores Nonaka e Takeuchi (1997) explicam também que o conteúdo criado por cada tipo de conversão é diferente, por exemplo: a socialização gera o conhecimento dito “compartilhado”, tal como os modelos mentais ou as habilidades técnicas compartilhadas, já a externalização gera o chamado “conhecimento conceitual” (metáforas e analogias), a combinação por sua vez gera o dito “conhecimento sistêmico”, através da geração de protótipos e tecnologias de novos componentes e, por fim, a internalização produz o “conhecimento operacional” ou a experiência prática de se trabalhar com determinado tipo de ofício.

Quando se pensa em grupos de estudo, o compartilhamento de conhecimento torna-se a principal meta a ser alcançada, ou seja, a sua socialização. Cada integrante contribui para a construção e reconstrução do conhecimento trazendo suas experiências e interagindo em prol do crescimento do grupo.

Porém, um dos principais problemas da gestão do conhecimento em grupos de estudos, concentra-se justamente na captura do conhecimento tácito do indivíduo e transferência deste conhecimento para terceiros (nova transformação em conhecimento tácito), pois se percebe que há uma dificuldade do ser humano em registrar/explicitar tudo aquilo que ele sabe fazer.

Polany (1996, p.4, *apud* NONAKA; TAKEUCHI, 1997) resume bem essa problemática na seguinte frase: “Podemos saber mais do que podemos dizer”. É neste momento que se torna propício questionar de que forma o processo de conversão do conhecimento tácito para tácito pode ser apoiado pelas tecnologias da informação.

Para tanto, a partir deste ponto, o presente capítulo fará uma discussão sobre o processo de socialização do conhecimento e suas principais características para posteriormente elencar e categorizar as principais ferramentas tecnológicas de apoio a este processo. O objetivo principal será permitir um melhor embasamento de requisitos funcionais para a proposta da ferramenta, que será desenvolvida no decorrer deste estudo.

2.3 O PROCESSO DE SOCIALIZAÇÃO DO CONHECIMENTO

Compartilhar saberes é uma das tarefas mais importantes apontadas pelos estudos na área de gestão do conhecimento, pois um saber isolado e “trancafiado” com uma pessoa ou um grupo restrito de pessoas tem seu valor reduzido frente àquele que circula, é difundido e é entendido por um maior número de pessoas.

Nonaka e Takeuchi (1997), afirmam que a base da criação do conhecimento organizacional é formada pelo conhecimento tácito mantido pelos indivíduos da organização e, portanto, o primeiro passo no processo de criação do conhecimento deve ser o compartilhamento do conhecimento tácito.

Vale ressaltar que esta não é uma tarefa simples, pois a transmissão de conhecimento tácito não se dá facilmente por meio de palavras, mas sim pela experiência.

O processo de socialização do conhecimento é definido, portanto, por esses autores como sendo um processo de compartilhamento de experiências e criação de conhecimento, tais como: modelos mentais (JOHNSON-LAIRD, 1983 *apud* NONAKA; TAKEUCHI, 1997, p.66) e habilidades técnicas compartilhadas.

Um exemplo corriqueiro de como se compartilha conhecimento tácito pode ser observado em academias de ensino de artes marciais, onde o aprendiz através da prática e da imitação do seu mestre aprende a aplicar movimentos com finalidades específicas de ataque ou defesa em uma luta corporal.

É importante perceber nesta situação que quanto mais o mestre conseguir compartilhar sua experiência prática, maiores serão as chances do aprendiz conseguir “absorver” e desenvolver tais habilidades. No mundo dos negócios uma analogia pode ser feita quando se pensa em aplicar treinamentos práticos para os integrantes de um departamento ou empresa.

Logo, o processo de socialização do conhecimento pode ser entendido como um processo que depende da experiência compartilhada (NONAKA; TAKEUCHI, 1997). Fialho, e outros (2006, p.111), complementam esta idéia lembrando que “a mera transferência de informações muitas vezes fará pouco sentido se estiver desligada das emoções associadas a elas e dos contextos específicos nos quais as experiências compartilhadas são embutidas”.

Cabe aqui uma reflexão sobre como a tecnologia poderia prover meios que auxiliassem o processo discutido neste tópico. Se o simples fato de armazenar e disponibilizar informações não promove a socialização do conhecimento, então o que pode ser considerado uma ferramenta/software de socialização do conhecimento?

2.4 TIPOS DE FERRAMENTAS PARA AUXÍLIO À SOCIALIZAÇÃO DO CONHECIMENTO

Conforme visto no tópico anterior, sistematizar a socialização do conhecimento não deve ser considerada uma tarefa trivial, pois depende essencialmente da capacidade de transmissão e absorção de experiências dos indivíduos envolvidos no processo.

Existem pesquisas sobre ferramentas para apoio aos processos envolvidos na Gestão do Conhecimento, porém poucas delas possuem o foco específico no processo de socialização, o qual é objeto de estudo desta dissertação.

Os autores Maier e Hädrich (apud JENNEX, 2008, p.541), por exemplo, conceituam os sistemas de gestão do conhecimento como “tecnologias para um gerenciamento do conhecimento efetivo e eficiente”, porém eles chamam a atenção para o fato de que o termo KMS (*Knowledge Management System*) tem sido freqüentemente utilizado de forma ambígua, como por exemplo, quando associado a ferramentas de gestão do conhecimento, a plataformas de gestão do conhecimento ou ainda ao uso de uma combinação de ferramentas tendo em mente a gestão do conhecimento.

O tema na área desperta opiniões divergentes e, nesta pesquisa, adotou-se a opinião dos trabalhos de: Wenger (2001), que traz um *survey* bastante referenciado por outros autores o qual contém uma proposta de categorização de ferramentas para apoio a comunidades de prática; Gallupe (2000), que também elaborou um *survey* que categoriza ferramentas para apoiar a gestão do conhecimento e Carvalho (2000) que propõe uma classificação dos diferentes tipos de ferramentas para a gestão do conhecimento.

Desta forma, serão adotadas como base principal para a proposta de categorização das ferramentas e elaboração deste capítulo, as idéias discutidas nestes três autores. Para isto serão identificadas as categorias em comum que mais

se aproximam da visão de socialização do conhecimento, bem como serão discutidas as suas principais características. Para complementação, outros autores servirão de apoio no aprofundamento dos conceitos envolvidos em cada categoria de ferramenta.

2.4.1 Portais de Conhecimento

Os portais de conhecimento têm sido amplamente adotados nas principais corporações, as quais vêm experimentando a idéia encarar seus empregados como “trabalhadores do conhecimento”. Esses sistemas presumem que tais trabalhadores participam de múltiplos grupos, comunidades, projetos e possuem a necessidade de gerenciar essas múltiplas “identidades” em sua vida profissional (WENGER, 2001).

Segundo Murray (apud DIAS, 2001, p. 7), esses portais unem as características de “portais de informações, cooperativos e de especialistas, sendo capaz de implementar tudo que os outros tipos de portais implementam e de fornecer conteúdo personalizado de acordo com a atividade de cada usuário”.

Segundo Wenger (2001), tais ferramentas contam com algumas características principais, tais como:

- a) Área de trabalho “customizável”;
- b) Gerenciamento de múltiplas visões sobre diversas fontes de informação importantes;
- c) Mecanismos de buscas;
- d) Espaços para conversações;
- e) Notificações e assinaturas;
- f) Capacidades de gerenciamento de projetos.

Um portal do conhecimento pode ser considerado uma ferramenta de socialização do conhecimento, quando este permite que as idéias sejam compartilhadas em um ambiente virtual único da organização e permite também a identificação de especialistas.

2.4.2 Sistemas de Trabalho em Grupo (*Groupware*)

Os sistemas de trabalho em grupo buscam reunir participantes de um mesmo projeto em prol da interatividade no meio virtual, bem como da integração dos processos e atividades realizadas dentro dos projetos.

Tais softwares constituem as chamadas ferramentas de *Groupware* e este termo é discutido e definido por Ellis, Gibbs e Rein (1991, p. 40) como sendo “sistemas baseados em computador que apóiam grupos de pessoas empenhados em uma tarefa (ou objetivo) em comum e que provêem uma interface para um ambiente compartilhado”. Wenger (2001), por sua vez, define que estes sistemas provêem um espaço online para um time de projeto conduzir o seu trabalho.

Gallupe (2000) destaca que as ferramentas de *groupware* possuem características para habilitar grupos a realizar tarefas como geração de idéias e busca de consenso em grupo.

Algumas características desejáveis para este tipo de sistema, segundo Wenger (2001), são:

- a) Planejamento, Controle e Compartilhamento de tarefas;
- b) Calendário da equipe;
- c) Notificação de eventos, mudanças e principais marcos de um projeto;
- d) Compartilhamento de documentos relacionados ao projeto;
- e) Painel de notícias;
- f) Painel de discussões;
- g) Controle de acesso dos membros;
- h) Mecanismos de busca.

O processo de socialização do conhecimento pode ser apoiado por sistemas de *groupware* no quesito compartilhamento de tarefas e idéias. Esse tipo de sistema promove a união de esforços das pessoas em prol de objetivos em comum (projeto), neste cenário os participantes precisam constantemente trocar experiências, o que promove, portanto, as transformações de conhecimento tácito para tácito entre eles.

2.4.3 Sistemas de Interações Síncronas

Os sistemas com troca de mensagens instantâneas constituem uma importante ferramenta para permitir a interação entre pessoas em tempo real a um baixo custo econômico.

Esse tipo de ferramenta já está bem difundido no mundo das tecnologias de informação e comunicação e permite a criação de situações de bate-papo on-line, reuniões, áudio e vídeo conferências entre pessoas geograficamente dispersas, dentre outras situações interessantes para quem está distante fisicamente ou não pode dispor do tempo presencial.

Algumas funcionalidades desejáveis para este tipo de ferramenta são apontadas por Wenger (2001):

- a) Áudio conferência;
- b) Vídeo conferência;
- c) Gravação de histórico de mensagens;
- d) Indicadores de reação de usuários (por exemplo: indicadores de humor);
- e) Controle de acesso;
- f) Chat.

O processo de socialização dentro de uma ferramenta deste tipo tende a ocorrer de forma mais rápida, pois existem aqui os conceitos de sincronia e instantaneidade das mensagens trocadas pelos participantes.

Com os recursos de áudio e vídeo, a troca de experiências pode ser enriquecida, por exemplo, com demonstrações visuais de execução de procedimentos em tempo real. Neste contexto, o aprendiz passa a ter a oportunidade de visualizar e ouvir o que um mestre tem a ensinar através das transmissões de movimentos e sons capturados pelas webcams e microfones instalados em seus computadores.

2.4.4 Ambientes para aprendizagem on-line

Os Ambientes Virtuais de Aprendizagem (AVA) são ferramentas que possibilitam a simulação de um ambiente de sala de aula através do computador. A

principal finalidade deste tipo de sistema é trazer elementos de interação *on-line* que possibilitem a construção de conhecimento por parte dos professores e alunos (CARLA NETTO apud FARIA, 2006).

Estes ambientes precisam da figura do mediador (geralmente o professor envolvido), o qual conduzirá o processo de aprendizagem através de postagens de artigos, gerenciamento das mensagens em fóruns, marcações de reuniões *on-line* (através de *chats*) e aplicação de testes/simulados no ambiente virtual.

Segundo Wenger (2001) esses sistemas geralmente possuem as seguintes características:

- a) Armazenamento de material de conteúdo;
- b) Formas de discussão de conteúdo abertas e diretas por parte dos estudantes;
- c) Processos de entregas de mensagens síncronas e assíncronas;
- d) Apresentações multimídias;
- e) Gravação e exibição de sessões de sala de aula para vários expectadores.

Um AVA pode contribuir para a socialização do conhecimento, pois esta modalidade de proposta pedagógica existente na educação à distância (EAD) possibilita a simulação de um ambiente de sala de aula interativa em computadores conectados à internet.

Este tipo de ferramenta faz uso de funcionalidades conhecidas, tais como: fóruns de discussão para comunicações assíncronas e *chats* para comunicações síncronas.

Os fóruns permitem ao estudante uma oportunidade maior de reflexão e aprofundamento do assunto discutido, pois as suas mensagens são postadas no momento que melhor lhe convier, diante de um tempo maior para sua elaboração (RAMOS apud FARIA, 2006), já os *chats* trazem melhor agilidade ao processo de troca de experiências, devido a seu caráter de troca de mensagens em tempo real.

Este assunto será novamente abordado no tópico sobre Ambientes Virtuais de Aprendizagem, já que dentre os sistemas de interação supracitados, este se configura como eixo de análise do presente trabalho.

Diante das ferramentas aqui pesquisadas, cabem ainda alguns questionamentos sobre o processo de socialização do conhecimento:

- a) Como identificar de forma rápida e precisa a(s) pessoa(s) que possuem o conhecimento sobre o assunto desejado para socializá-lo com os demais componentes do grupo?
- b) De que forma é possível entrar em contato com esta(s) pessoa(s)?

2.5 BUSCA POR ESPECIALISTAS

O mundo atual é marcado por um fenômeno muito discutido por pensadores contemporâneos: a hiper-especialização. Tal fenômeno se percebe quando é possível observar a crescente demanda de pessoas que têm buscado uma formação continuada, no intuito de se inserir em um mercado de trabalho cada vez mais competitivo, especializando-se, assim, em determinadas áreas.

Nesta direção, os especialistas surgem como importantes atores sociais, que representam as áreas do conhecimento e são porta-vozes das inovações científicas na sociedade.

Porém, um problema antagônico que pode ser percebido é justamente a dificuldade em se encontrar pessoas que são especialistas em assuntos a tempo de estas poderem ajudar a encontrar soluções diversas em suas áreas respectivas de conhecimento.

Em equipes com uma quantidade grande de pessoas, por exemplo, nem sempre é possível conhecer as características, competências e habilidades de cada um de seus integrantes. O dinamismo do trabalho dessas pessoas pode depender de uma dica ou esclarecimento sobre um determinado assunto, o qual pode ser de domínio de um ou mais especialistas que ajudam ou compõem a própria equipe.

Os autores Mattox, Maybury e Morey (1999) afirmam que muitos grupos sociais já se depararam com questões importantes, que poderiam ser respondidas se a pessoa certa para se perguntar pudesse ser encontrada a tempo de formular uma resposta.

Os referidos autores lembram ainda que no mundo empresarial a distribuição das equipes, a diminuição do tamanho dos projetos e a pressão por custo e tempo,

levam os envolvidos no processo a uma necessidade de captar opiniões especializadas com eficiência. Em outras palavras, estes buscam descobrir rapidamente quem sabe o quê, para encontrar e formar com agilidade equipes de especialistas.

Segundo Santos e Salvador (2009):

O processo de busca por especialistas tem sido alvo de muitos pesquisadores, os quais analisam possibilidades de automatização do mesmo por meios computacionais. Tais esforços em pesquisas e obras acadêmicas acabaram por criar o termo *Expert Find* (ou Busca de Especialista). (SANTOS; SALVADOR, 2009, p.2).

Os autores supracitados complementam ainda a definição deste termo destacando a clareza de Balog e Rijke (2007 *apud* Santos; Salvador, 2009 p.2), ao explicarem o termo *Expert Find* da seguinte forma: “A busca por especialistas se destina à tarefa de encontrar a pessoa certa com as habilidades e conhecimento apropriados: ‘Quem são os especialistas no tópico X?’”.

Neste contexto, é importante destacar que a busca por especialistas pode ser uma atividade puramente manual, tal como a análise de autoria de documentos, *e-mails* ou da participação de pessoas em uma rede de relacionamento, porém é interessante que se tenha uma forma de automatizar tal processo. Surge, portanto, a idéia de se criar ferramentas para apoio à busca de especialistas, as chamadas *Expert Finders*.

Uma ferramenta do tipo *Expert Finder* (ou buscador de especialistas), constitui-se basicamente de uma máquina que promove a busca por pessoas especialistas em determinados assuntos. Tal máquina geralmente levará em consideração aspectos como: quantidade de publicações de uma pessoa sobre o(s) assunto(s) desejado(s); quantidade de referências feitas por terceiros a esses autores; número de participações de uma pessoa em uma comunidade; dentre outros (BECERRA-FERNANDEZ, 2006).

Muitos estudos têm sido desenvolvidos no sentido de implementar ferramentas de busca por especialistas e, neste contexto, faz-se necessário entender quais são os potenciais contextos de aplicação deste tipo de sistema.

Yimam-Seid e Kobsa (2003) fazem uma discussão interessante na qual destacam que um buscador de especialistas tem seu potencial expandido quando associado a outros sistemas de informação empresarial, tais como: sistemas de

gestão do conhecimento, sistemas colaborativos (CSCW) ou mais conhecidos como sistemas de *Groupware*, sistemas do tipo “recomendadores” (ou sistemas de filtragem colaborativos), e ainda os sistemas de mercado eletrônico para especialidade/competência humana.

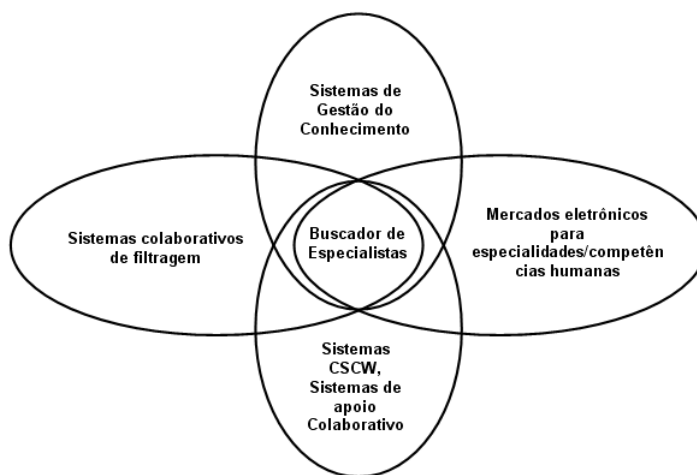


Figura 2 - Buscadores de especialistas integrados a sistemas de informações organizacionais

Nota: Adaptado de Yimam-Seid e Kobsa (2003).

É importante ressaltar que o presente trabalho científico possui uma abordagem similar à defendida por esses autores, apesar de se perceber um foco claramente voltado ao âmbito empresarial por parte destes, pois visa integrar uma solução de busca por especialistas a um ambiente virtual de aprendizagem, que neste contexto pode ser entendido como um tipo de sistema colaborativo com finalidade de servir ao mundo acadêmico.

A literatura mostra que muitas ferramentas de busca por especialistas vêm sido desenvolvidas ao longo do tempo. Segundo Jung, e outros, (2007, p. 56) “as fontes para encontrar especialistas são diversas: documentos, programas, *e-mails*, bases de dados, citações, comunidades, dentre outras”. Maybury (2006) por sua vez, complementa a idéia dos autores supracitados destacando que as fontes podem ser também compostas de auto-declarações, resumos e páginas *web*.

Ainda segundo Maybury (2006), ferramentas deste tipo podem ser analisadas levando-se em consideração, não só os seus tipos de fontes, mas também pelo:

- a) Tipo de processamento que o software aplica: tais como ranking automático de especialistas, extração de entidades a partir do texto (pessoas, localidades, organizações), análise de rede social (determinar relacionamentos entre os especialistas), suporte a línguas estrangeiras ao inglês e, por fim, identificação de autoria (publicações de documentos).
- b) Tipo de consulta suportado pela ferramenta: incluindo palavras-chave, consulta booleana, linguagem natural, ou exibição taxonômica do conteúdo.
- c) A forma como os resultados são exibidos: lista ordenada de especialistas, lista de documentos ou artefatos produzidos ou utilizados por especialistas, e conceitos relacionados ao tópico de especialidade que o usuário está buscando.
- d) Propriedades do sistema: tipos de suporte a interoperabilidade provido pelo sistema, está em uso operacional de modo geral, ou ainda, os tipos de suporte à privacidade providos pela ferramenta.

Alguns exemplos de ferramentas de busca por especialistas são destacados em Maybury (2006) e serão aqui replicados (vide figura 3) com o intuito de uma comparação futura com a solução proposta no presente trabalho. Vale ressaltar que estes exemplos são todos de soluções disponíveis comercialmente, nenhuma delas se enquadra como software livre ou de código fonte aberto.

Ferramentas de Busca por Especialistas

Capacidade	Ferramentas de Busca por Especialistas																						
	Fontes					Processamento				Busca			Resultados		Sistema								
Total	auto declaração	e-mail	documentos	resumos	paginas web	bancos de dados	ranking	extração de entidades	análise de rede social	língua estrangeira	identificação de autoria	palavra-chave	booleana	linguagem natural	taxonomia (exibição)	lista de especialistas	documentos relacionados	conceitos relacionados	interoperabilidade	em uso operacional	privacidade		
Parcial																							
Nenhuma																							
Produto																							
TACIT										2													
AskMe																							
Autonomy										70													
Endeca										250													
Recommind										200													
Trivium										6+													
Entopia										6													

Figura 3 - Sistemas de Busca por Especialistas

Nota: Adaptado de Maybury (2006).

Conforme pode ser observado na figura acima, a maioria das ferramentas listadas possui amplo suporte a diversidade de fontes de busca por especialistas, porém divergem muito quanto ao suporte que é provido aos tipos de processamento, sendo que neste contexto é válido destacar que todas elas possuem *ranking* de especialistas e nenhuma possui suporte a identificação de autoria. É possível observar também que todas em algum nível possuem suporte a línguas estrangeiras ao inglês, com destaque para a Endeca com suporte a 250 línguas diferentes.

Ainda analisando a figura 3, é possível verificar que os tipos de consulta por palavra-chave e booleano predominam entre as ferramentas com destaque para as ferramentas AskMe, Autonomy e Endeca que dão suporte a todos os tipos de consulta.

Quanto à forma de exibição dos resultados, a unanimidade está na listagem de especialistas, destaque para a ferramenta Entopia que dá amplo suporte a todos os tipos de exibição.

Finalmente, em relação às propriedades de sistema, todas as ferramentas analisadas pelo autor possuem amplo suporte à interoperabilidade e estão em uso operacional.

2.6 CONSIDERAÇÕES FINAIS

A gestão do conhecimento é uma área com grande demanda por pesquisas relativas às tecnologias de informação e comunicação. O conhecimento passou a ser o centro das atenções das instituições, que agora se preocupam com as melhores formas de captar e disseminar este conhecimento para seus diversos setores.

Este capítulo discutiu os dois tipos básicos de conhecimento (tácito e explícito) e sua importância no estudo sobre gestão do conhecimento. Segundo os autores Nonaka e Takeuchi (1997), o conhecimento passa por algumas transformações, quais sejam: a socialização (do tácito ao tácito), a externalização (do tácito ao explícito), a combinação (do explícito ao explícito) e a internalização (do explícito ao tácito).

Porém o conhecimento pode ser considerado, em muitos casos, uma matéria prima de difícil gerenciamento, pois no centro desta problemática estão as pessoas, as quais podem contribuir, mas também criar barreiras para a disseminação do conhecimento na organização. Tais pessoas são peças fundamentais, pois parte delas a capacidade criativa e inovadora no dia a dia de uma instituição.

A busca por especialistas dentro das organizações virou, portanto, tema de discussão para diversos autores e, no presente trabalho, buscou-se mais uma vez analisar uma forma em que a tecnologia possa apoiar este processo, principalmente o que envolve a socialização do conhecimento de especialistas. Para tanto, algumas ferramentas de auxílio à socialização do conhecimento foram analisadas, bem como as ferramentas de busca por especialistas, visando criar um embasamento teórico para solução proposta nesta pesquisa, a qual será discutida nos capítulos 5 e 6 do presente texto.

O próximo capítulo fará uma análise sobre alguns dos principais ambientes virtuais de aprendizagem, de forma a servir como base para se pensar a funcionalidade da solução aqui proposta.

3 AMBIENTES VIRTUAIS DE APRENDIZAGEM

O aprofundamento acerca do conceito Ambientes Virtuais de Aprendizagem dá continuidade a discussão iniciada no capítulo anterior e pretende abordar alguns exemplos de ferramentas de apoio à educação em grupos virtuais (à distância), comumente chamadas de LMS (*Learning Management Systems* – Sistemas de Gerenciamento de Aprendizagem), como forma de encontrar um ponto de partida para a definição dos requisitos funcionais da solução proposta.

Durante o processo de elaboração desta dissertação foram encontradas as seguintes ferramentas com proposta educacional e com grande utilização no mercado: AulaNET, Blackboard, Moodle, TelEduc e Dokeos.

Pretende-se explicar quais são as principais funcionalidades destes ambientes, bem como analisar os aspectos referentes à tecnologia sob a qual foram desenvolvidos, as versões nas quais se encontram atualmente, suas formas de extensão por parte de terceiros e qual a estratégia utilizada para acompanhar o nível de interação dos participantes ao longo de cada curso.

3.1 AULANET

Esta ferramenta foi criada pelo Laboratório de Engenharia de Software (LES) da Universidade Católica do Rio de Janeiro (PUC-Rio), com a finalidade principal de servir como um ambiente virtual de ensino-aprendizagem para cursos oferecidos na modalidade EAD. O AulaNet se baseia a abordagem da teoria do *groupware*, ou seja, trata-se de um sistema para suporte ao trabalho em grupos onde seus integrantes se encontram geograficamente dispersos (GEROSA; RAPOSO; FUKS; LUCENA, 2004).

O AulaNet é uma ferramenta *freeware* (seu *download* e instalação são gratuitos), que pode ser encontrada nas versões em português, inglês e espanhol. Ela pode ser melhor entendida, observando-se os módulos que a compõem na figura 4.

Conforme pode-se observar na figura4, o ambiente AulaNet se baseia atualmente no modelo 3C proposto por Ellis, Gibbiss e Rein (1991) os quais propõem que ferramentas de *groupware* devem prover 3 tipos de serviços principais,

sendo eles: o serviço de comunicação, de cooperação e de coordenação entre os indivíduos participantes de cada grupo.



Figura 4 - Modelo de serviços do AulaNet classificados em função do modelo 3C

Fonte: Gerosa e outros (2004)

A ferramenta possui ainda uma proposta de arquitetura extensível baseada na idéia de *component frameworks*, os quais permitem que o ambiente possa ser estendido através da criação e conexão de novos componentes ao seu framework central. Segundo os autores Gerosa e outros (2004, p.2), a arquitetura do AulaNet possui dois níveis de componentização. Os componentes principais são os serviços oferecidos pelo ambiente, tais como bate-papo, fórum, agenda, lista de discussão, dentre outros. Tais serviços podem ser “plugados e desplugados de forma a montar um ambiente personalizado para cada situação”.

Os autores complementam ainda informando que há um segundo nível de componentes no AulaNet o qual acontece dentro dos serviços anteriormente citados, ou seja, cada serviço existente também se utiliza de conceitos de componentes, conforme pode ser observado a seguir na figura 5:

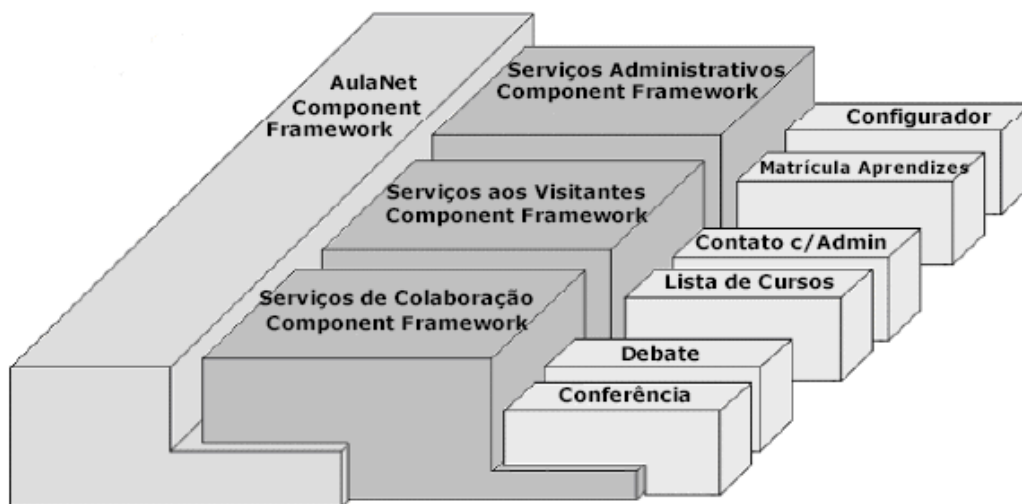


Figura 5 - Parte da arquitetura do ambiente AulaNet

Fonte: Gerosa e outros (2004, p.2).

Vale ressaltar que o ambiente foi desenvolvido utilizando a linguagem Java, que é uma linguagem amplamente difundida entre os desenvolvedores espalhados pelo mundo, e a sua arquitetura base de *software* se utiliza da plataforma J2EE¹, ambas, linguagem e plataforma, fazem parte dos produtos da empresa Sun Microsystems (EDUWEB, 2008).

Durante esta pesquisa foi possível perceber uma série de trabalhos acadêmicos voltados para o ambiente AulaNet que trazem contribuições significativas em cima desta ferramenta. Percebe-se, por exemplo, que há uma flexibilidade notória no quesito extensibilidade deste *software*, vide o trabalho de Filippo, Fuks e Lucena (2005), os quais propuseram uma extensão do serviço de conferências contido no AulaNet para que este possa ser executado em dispositivos móveis de diversos tipos (celulares, *Personal Digital Assitant*² - PDA), a esta extensão eles deram o nome de AulaNetM.

Segundo os autores, foram feitas novas interfaces das telas de conferências do AulaNet para serem executadas em PDAs inicialmente, pois estes oferecem uma tela maior para a leitura das mensagens de conferências frente aos celulares, propiciando ao usuário portanto um melhor conforto visual.

¹ J2EE (Java to Enterprise Edition): uma arquitetura padrão para desenvolvimento de aplicações corporativas que utilizam a linguagem Java (SUN, 2008)

² Este termo se refere a dispositivos móveis, como por exemplo, *palm tops* e computadores de bolso.

O AulaNet possui uma funcionalidade para acompanhamento das interações feitas pelos estudantes no sistema intitulada Relatórios de Participação, porém esta é disponibilizada apenas para o perfil Coordenador. Neste caso os professores não podem ser notificados ou verificar o nível de participação dos estudantes ao longo do curso, nem tem condições de identificar os estudantes mais ativos através de recursos da ferramenta pré-instalados.

Os relatórios possíveis para coordenadores são: conceito médio dos participantes em todos os serviços da ferramenta, porcentagem média de contribuição efetiva em todos os serviços, contato com docentes, lista de discussão, conferência, debates, tarefas, avaliação e co-autoria de aprendiz.

3.2 MOODLE

O Moodle é um ambiente virtual de ensino-aprendizagem de código fonte aberto (*OpenSource*), o que o caracteriza portanto como sendo uma ferramenta com liberdade de uso, distribuição e modificação, por parte de terceiros.

Ele foi desenvolvido na década de 1990, pelo então *webmaster* Martin Dougiamas (atualmente líder do projeto), o qual era um administrador de sistemas e cuidava do WebCT (outra ferramenta de EAD) da *Curtin University of Technology* (MOODLE, 2008a).

A palavra Moodle foi inicialmente pensada como sendo um acrônimo para *Modular Object-Oriented Dynamic Learning Environment* (Ambiente de Aprendizado Dinâmico Modular Orientado a Objeto), o que traduz em poucas palavras para quem serve e como foi projetado o sistema.

Existem muitas instituições que já se utilizam desta ferramenta de EAD, somente no Brasil, mais de 2900 sites estão registrados no site Moodle.org. Os dados gerais coletados no *site* do Moodle em novembro de 2008, revelam que existem 45916 *sites* de 199 países registrados e a grande maioria destes pode ser consultada a partir do próprio *site* (MOODLE, 2008b).

Percebe-se que o Moodle vem se consolidando com uma ampla utilização em todo o mundo, o que provavelmente pode ser atribuído ao fato deste sistema ser disponibilizado através da filosofia do *software* livre (licença GPL – GNU *General Public License*).

A arquitetura da ferramenta se baseia em uma interface *web* escrita na linguagem PHP contendo uma divisão de suas funcionalidades por módulos, sendo eles o de Administração do *Site*, Administração de Usuários, Administração de Curso, Tarefa, *Chat*, Pesquisa de Opinião, Fórum, Questionário, Recursos, Pesquisa de Avaliação e Laboratório de Avaliação. Outros módulos podem ser desenvolvidos para a ferramenta, seguindo-se a documentação para desenvolvedores disponibilizada no site.

Uma pessoa com habilidades de desenvolvimento de sistemas na tecnologia envolvida tem a liberdade e a possibilidade de desenvolver novos módulos e/ou *plugins* para o Moodle. Estes módulos podem ser posteriormente incorporados à ferramenta mediante aprovação prévia dos responsáveis pelo projeto.

Existe uma vasta coleção de módulos e *plugins* que podem ser baixados no *site* da ferramenta e não é o foco desta pesquisa listar todas essas possibilidades, mas é importante destacar que alguns deles foram desenvolvidos visando extrair informações sobre o nível de interação dos estudantes ao longo do curso, tais como: o módulo de exibição dos usuários mais ativos de um determinado curso e o módulo que trata da estimativa de tempo dedicado ao curso por cada um de seus participantes (MOODLE, 2008c)

3.3 BLACKBOARD

O Blackboard ou quadro-negro, traduzindo para o português, é o terceiro ambiente virtual de ensino-aprendizagem a ser analisado nesta pesquisa. Trata-se de um sistema de licença proprietária, pertencente à empresa Blackboard fundada em 1997 e que conta com uma suíte acadêmica que pode ser integrada aos sistemas acadêmicos das instituições de ensino (BLACKBOARD, 2008).

O sistema não se baseia em uma filosofia de educação específica, pois seu projeto tem como objetivo maior o foco no incremento das realizações do estudante. Segundo o próprio site da ferramenta um “instrutor” pode usar um método como o aprendizado baseado em problemas ou o de estudos de caso que a ferramenta terá condições de viabilizar tais necessidades, porém o projeto considera três áreas chave para definir suas funcionalidades, são elas: Instrução, Comunicação e Avaliação.

Atualmente o Blackboard possui uma versão chamada Academic Suíte a qual compreende o Blackboard Learning System, o Blackboard Content System e o Blackboard Community System (BLACKBOARD, 2008). No contexto destes subsistemas, o mais adequado para esta pesquisa será o Learning System, tendo em vista as características das suas funcionalidades. Um dos objetivos do estudo é traçar uma comparação entre esta e as outras ferramentas de EAD aqui abordadas, considerando os seus aspectos mais pertinentes.

O Learning System pode ter suas funções separadas nas três áreas citadas anteriormente, como se segue:

a) Instrução

- Course Management – criação e gerenciamento dos cursos
- Content Authoring – interface de edição de textos similar ao processador Word da Microsoft
- Adaptative Release – elaboração (dos professores) de caminhos de aprendizagem para os estudantes, (determinando quando estes podem acessar os itens, as discussões ou outras atividades de aprendizagem)
- Syllabus Builder – criação/upload de planos de ensino
- Learning Units – permite a criação de lições seqüenciadas e o controle de navegação dos estudantes através destas
- Teaching and Learning Tools – disponibiliza ferramentas para professores e estudantes tais como criação de glossário e o quadro negro elétrico
- Professional Information Management – disponibiliza ferramentas para que discentes e docentes gerenciem o dia-a-dia dos seus trabalhos tais como calendário, tarefas e mensagens

b) Comunicação

- Discussion Board – uma espécie de fórum para permitir discussões assíncronas entre os participantes dos cursos. Tais fóruns podem estar embutidos em determinadas áreas de conteúdo
- Group Projects – permite a criação de múltiplos grupos de estudantes. Cada grupo pode ter sua própria área de troca de

- arquivos, Discussion Board, Virtual Classroom (a ser discutido posteriormente) e e-mail de grupos
 - Virtual Classroom – suporta interação síncrona (ao vivo) para simulação de uma sala de aula virtual com possibilidades de texto, áudio e vídeo
- c) Avaliação
- Assessments and Surveys – permite que os professores criem avaliações e exames automaticamente pontuados.
 - Gradebook – armazena os resultados de performance dos estudantes.
 - Assignments – permite que os professores criem itens de tarefas para que os estudantes submetam suas respostas.

A ferramenta possui ainda possibilidades de extensão de suas funcionalidades através de uma tecnologia intitulada *Building blocks*. Trata-se de um kit para desenvolvedores interessados em criar soluções e adicioná-las ao sistema. Este kit é disponibilizado gratuitamente e contém a API (Interface de Programação de Aplicação) do Blackboard.

Esta pesquisa não obteve acesso ao Blackboard para instalação e testes devido ao fato do mesmo ter um custo associado para que fosse disponibilizado, também não havia uma versão de testes gratuita disponível no site, o que inviabilizou uma análise da estratégia de acompanhamento das interações dos estudantes no sistema.

3.4 TELEDUC

O TelEduc, assim como o AulaNet, também é um ambiente para realização de cursos à distância de autoria brasileira e que, neste caso, se encontra em desenvolvimento pelo Núcleo de Informática Aplicada a Educação (NIED) da Universidade Estadual de Campinas (UNICAMP) e seu projeto está sob a orientação de Professora Heloísa Vieira da Rocha (TELEDUC, 2008).

Ainda segundo o site da ferramenta, o ambiente tem sido utilizado para viabilizar cursos à distância nesta instituição desde 1998 e vem sendo acompanhado e avaliado pelo NIED durante este tempo.

Este ambiente também é disponibilizado através da licença GPL, versão 2, trazendo consigo, portanto, todas as implicações já vistas anteriormente em outros ambientes que possuem o mesmo tipo de licença.

A estrutura do ambiente conta com recursos distribuídos de acordo com o perfil de seus usuários, os quais neste caso podem ser: alunos ou formadores (TELEDUC, 2008).

Dentre os recursos disponíveis, alguns são bem comuns e já foram mencionados nas outras ferramentas deste tipo, tais como: agenda, a lista de avaliações, lista de atividades a serem realizadas, fóruns de discussão e o recurso de bate-papo. Há ainda outros recursos que podem ser destacados como:

- a) Mural – espaço reservado para os participantes disponibilizarem informações consideradas relevantes para o contexto do curso.
- b) Diário de Bordo – um espaço para o relato de experiências (sucessos, dificuldades, dúvidas) ao longo do curso, as quais podem ser compartilhadas ou não com os demais participantes.
- c) Acessos – acompanhamento de frequência de acesso dos usuários ao curso e às suas ferramentas.
- d) Intermap (disponível apenas para o perfil dos formadores) – permite visualizar a interação dos participantes do curso nas ferramentas: correio, fóruns de discussão e bate-papo.

Não foram encontradas informações no *site* do projeto que identifiquem se o sistema possui formas de extensão das suas funcionalidades por parte de terceiros. Para efeito de análise, na tabela comparativa ao final deste capítulo a ferramenta será marcada como N/A (Não se aplica).

É importante destacar que a ferramenta possui um recurso chamado Intermap, o qual permite a exibição de um mapeamento das interações dos estudantes em forma de grafo ou tabela contendo informações sobre as quantidades de mensagens trocadas por eles ao longo do tempo nas funções de bate-papo, fórum e correio.

O grafo do Intermap marca quem interagiu com quem criando uma linha de ligação entre eles, sendo que a ferramenta destaca a cor dos docentes dentre os demais participantes para facilitar a visualização de interações feitas com os professores. Os participantes do grupo que não tiverem interagido no sistema até o momento da visualização aparecerão no grafo sem ligação com nenhum outro.

3.5 DOKEOS

O Dokeos é o último sistema LMS a ser abordado nesta pesquisa, o qual também é disponibilizado sob a licença GPL, versão 2 e possui características bem parecidas com os outros sistemas já abordados anteriormente.

O sistema possui uma versão de testes no seu site o que facilitou a análise em termos de usabilidade da ferramenta e das funcionalidades disponibilizadas. Dentre estas funcionalidades, observa-se que o lado negativo do sistema é que, pelo menos na versão de testes disponibilizada no site, o sistema não possui uma ferramenta de edição de textos com possibilidade de formatação, ou seja, o usuário precisa se apropriar de conhecimentos de HTML (*HyperText Markup Language*) para digitar mensagens nos fóruns e configurar as informações dos cursos com formatações especiais de letras.

Tal qual o Moodle, esta ferramenta também foi feita utilizando a tecnologia PHP, porém a diferença é que ela possui duas formas de extensão aparentemente bem simplificadas, são elas: os Plugins e as Ferramentas (também chamadas de Módulos).

No projeto do sistema, os plugins são pequenos componentes que podem ser acoplados em determinadas áreas das telas do Dokeos, já os módulos são *scripts* mais completos que adicionam funcionalidade a todos ou determinados cursos conforme a necessidade do desenvolvedor (DOKEOS, 2008).

A simplicidade da criação destas extensões pode ser observada, pois é necessário apenas que o desenvolvedor armazene seus arquivos em subpastas determinadas do diretório principal onde o sistema está instalado e altere alguns arquivos indicando a existência deste novo item (seja ele um *plugin* ou uma ferramenta/módulo).

Minha lista de cursos | Meu espaço | Logout (mlapa)

Desenvolvimento de Games > Acompanhamento > Detalhes do estudante no curso

Imprimir | Exportar como arquivo CSV

Informações	Acompanhamento	Ação
Nome : ██████████	First connection 10 Janeiro, 2009	Enviar e-mail
E-mail : ██████████	Última conexão 12 Janeiro, 2009	Access details
Tel. Sem telefone	Tempo gasto na plataforma 0:04:53	
Online : Não	Progresso 0 %	
	Pontuação 0 %	

Desenvolvimento de Games | Número total de conexões para esse curso : 4 | Tutor : ██████████

Rotas de aprendizado	Tempo	Pontuação	Progresso	Última Conexão	Detalhes
Sem rota de aprendizado					
Exercícios	Pontuação	Tentativas	Corrigir este teste		
Exercício de exemplo		0 %	0		
Teste		0 %	0		
Outras ferramentas					
Publicações de Estudantes	0				
Mensagens	1				
Links acessados pelo usuário	2				
Documentos baixados pelo usuário	0				

Figura 6 – Exemplo de tela de acompanhamento do Dokeos

Um dos pontos fortes desta ferramenta é a funcionalidade de acompanhamento dos estudantes, conforme pode ser observado na figura 6, a qual possui informações sobre a quantidade de mensagens enviadas pelos estudantes, o tempo gasto por estes na interação com sistema, bem como seus progressos em avaliações e nas rotas de aprendizagem do sistema.

3.6 COMPARATIVO ENTRE OS AMBIENTES PESQUISADOS

Existe uma série de AVA no mercado, alguns são de código aberto e outros com licença proprietária, e o presente trabalho procurou analisar cinco ferramentas encontradas com grande aceitação na comunidade acadêmica, analisando-as à luz de cinco aspectos, a fim de proporcionar uma visão comparativa para futuros usuários e pessoas interessadas em contribuir com o avanço deste tipo de ferramenta.

Os aspectos a serem analisados são:

- Tecnologia – que linguagens ou plataformas de desenvolvimento serviram de suporte para o desenvolvimento da ferramenta
- Tipo de Licença – indicação se os direitos de uso e desenvolvimento estão atrelados a algum fabricante ou se constitui uma ferramenta de código aberto

- c) Possibilidades de expansão – apresenta a(s) forma(s) de se estender a capacidade funcional da ferramenta
- d) Estratégias para verificar o nível de participação dos envolvidos – descrição de como a ferramenta pode prover formas de verificação e análise sobre a participação das pessoas envolvidas no curso
- e) Analisa o conteúdo das mensagens – indicação da existência de algum tipo de análise sobre o texto das mensagens postadas pelos participantes dos cursos, por parte da ferramenta

O quadro abaixo resume e enquadra ambientes virtuais de aprendizagem pesquisados diante dos aspectos anteriormente citados.

Ambiente	Tecnologia	Tipo de Licença	Possibilidades de Expansão	Estratégias para verificar o nível de participação dos envolvidos	Analisa o conteúdo das postagens?
AulaNet	Java/J2EE	GPL (Free)	Framework	Relatórios para coordenadores: conceito médio dos participantes nos serviços; porcentagem média de contribuição efetiva nos serviços; contato com docentes; listas de discussão; conferência; debates; tarefas; avaliação e co-autoria de aprendiz	Não
Moodle	PHP	GPL (Free)	Módulos e Plugins	Esforços de terceiros na criação de módulos para: exibição dos usuários mais ativos e para estimativa de tempo dedicado ao curso pelo participante	Não
Blackboard	Não foi possível analisar	Proprietária	Buildingblocks	Não foi possível analisar tal aspecto da ferramenta	Não foi possível analisar
TelEduc	PHP	GPL (Free)	N/A	Recurso Intermap: mapeamento das interações entre os participantes em forma de grafos ou tabelas.	Não
Dokeos	PHP	GPL (Free)	Módulos e Plugins	informações sobre a quantidade mensagens enviadas, o tempo gasto na interação com sistema, bem como os progressos em avaliações e nas rotas de aprendizagem.	Não

Quadro 1 - Comparativo de ambientes virtuais de aprendizagem

A partir da categorização acima apresentada é possível observar quais são as principais características de cada ferramenta abordada, bem como analisar suas semelhanças e diferenças.

Pode-se perceber inicialmente que três ambientes foram feitos utilizando a linguagem PHP e que o AulaNet foi feito usando a plataforma Java. A tecnologia utilizada no Blackboard não pôde ser analisada, devido ao fato de que este é o único sistema, dentre os pesquisados, que possui licença proprietária. Todos os outros contam com a filosofia do software livre para se desenvolver e evoluir suas funcionalidades.

Ao se analisar as possibilidades de expansão, percebe-se que o TelEduc é o único sistema que não possui uma forma de extensão de suas funcionalidades por parte de terceiros documentada em seu site.

A maior parte dos ambientes se propõe a exibir relatórios com quantidades de postagens e interações dos participantes, sendo que o TelEduc possui uma forma diferenciada de visualização deste tipo de informação, traduzida na capacidade de gerar grafos de interação entre os docentes e discentes dos cursos.

Por fim, o quadro revela um fato interessante: nenhum destes ambientes se propõe a fazer uma análise sobre o texto das mensagens postadas por seus participantes, para acompanhar o nível de interação entre eles. Este acompanhamento se dá através da quantidade de informações ou interações e não se faz reflexão alguma sobre o conteúdo destas postagens.

3.7 UM OLHAR PROPOSITIVO SOBRE AMBIENTES VIRTUAIS DE APRENDIZAGEM

Os estudos no âmbito da gestão do conhecimento apontam para a necessidade de se mapear o conhecimento dos indivíduos que compõem as organizações, de forma que se ganhe agilidade na busca de pessoas aptas a resolver problemas mais específicos.

O filósofo Pierre Lévy (1999) em sua obra traz a idéia de se criar árvores do conhecimento (ou das competências) dos componentes de uma comunidade como forma de catalogar suas habilidades comportamentais e os conhecimentos teóricos. A idéia básica é valorizar as pessoas hierarquicamente em um grupo com base em suas competências, conforme se pode perceber no trecho abaixo:

A representação em árvore de conhecimentos permite localização, por simples inspeção, da posição ocupada por determinado saber em um momento dado e os itinerários de aprendizagem possíveis para ter acesso a esta ou aquela competência. Cada indivíduo possui uma imagem pessoal (uma distribuição original de brevês) na árvore, imagem que ele pode consultar a qualquer momento. Chamamos essa imagem de 'brasão' da pessoa, para marcar que a verdadeira nobreza de nossos dias é conferida pela competência. (LÉVY, 1999, p.178).

Portanto a idéia das árvores do conhecimento de Lévy pode ser entendida como uma ferramenta para Gestão do Conhecimento, pois ela apresenta uma forma de se mapear as pessoas pelas suas competências e habilidades, onde, através do formato de árvore que é aplicado ao seu mapeamento, é possível também chegar a conclusões a respeito de que pessoas seriam especialistas em determinados assuntos.

O conceito de busca por especialistas, discutido no presente trabalho trata-se de uma idéia que possui um objetivo bem semelhante ao das Árvores do Conhecimento de Lévy.

Em algumas situações em grupos de pesquisa ou aprendizagem virtual, torna-se igualmente importante saber quem está de fato contribuindo para o bom andamento da pesquisa ou do curso. Esta informação serve para que professores/coordenadores de equipe possam, por exemplo, tomar decisões sobre lideranças e formações de subgrupos.

A idéia é verificar a participação efetiva de cada componente dentro do grupo e uma das formas, portanto, poderia ser a utilização de uma ferramenta do tipo busca de especialistas para analisar as postagens sobre um assunto específico.

Conforme visto anteriormente, na análise feita sobre os ambientes virtuais de aprendizagem, percebe-se que nenhum deles trouxe consigo uma funcionalidade que permitisse uma avaliação do conteúdo das mensagens postadas nos grupos virtuais, a fim de se perceber a pertinência do mesmo frente ao assunto a ser discutido e o nível de importância da mensagem atribuído pelos demais participantes interessados.

O presente trabalho visa justamente criar uma proposta (a qual será detalhada no capítulo 5 do presente trabalho), para suprir esta lacuna dos ambientes virtuais de aprendizagem, uma vez que ela se propõe a combinar algumas técnicas

de busca de especialistas para serem executadas nas bases de informações geradas por um ambiente virtual de aprendizagem dentre os pesquisados.

4 MINERAÇÃO DE TEXTOS

A necessidade de localizar, categorizar e analisar conceitos em meio à quantidade de dados e informações produzidas no mundo contemporâneo tomou grandes proporções com o advento da globalização e com o suporte cada vez mais intenso das Tecnologias de Informação e Comunicação.

As questões acima orientam o olhar para a seguinte constatação: a Contemporaneidade rompe com uma visão positivista acerca do processo de produção e aquisição do conhecimento, colocando os sujeitos em uma relação de busca por saberes, que estão em permanente transformação (MORIN, 2001).

Busca-se, para fundamentar este estudo, o campo de análise formado por sujeitos inseridos na relação ensino-aprendizagem, em espaços virtuais de (re)elaboração do saber, em outras palavras, atores da área de educação que interagem em fóruns de discussão acadêmico.

Segundo os autores Nonaka e Takeuchi (1997), o conhecimento produzido é, por vezes, explorado e aproveitado de forma inadequada, pois, em meio às discussões e trocas de informação, muito ruído (trocas de mensagens inúteis ao contexto de um grupo de discussão, por exemplo) pode ser gerado e informações valiosas podem ser desperdiçadas.

Quando se pensa em armazenamento de informações, percebe-se que diversas tecnologias evoluíram nesse sentido, proporcionando um cenário cada vez mais complexo, onde o surgimento de diversos tipos de *softwares* (Sistemas de Informação, Bancos de Dados, Aplicativos de Escritório...), infra-estruturas de comunicação em rede e dispositivos (Servidores, Computadores Pessoais, Notebooks, *Palms*, Celulares...) ampliou a capacidade de comunicação das pessoas e de produção de grandes volumes de dados, ao longo do tempo (RESENDE e outros apud RESENDE, 2005). Ainda segundo esses autores, muitos desses dados podem constituir informações valiosas, mas eles pouco valem, quando não é possível compreender o seu significado/importância no contexto.

4.1 MINERAÇÃO DE DADOS E KDD

A Mineração de Dados (MD) é uma área de pesquisa que analisa justamente as técnicas e formas de descobrir informações úteis em meio ao emaranhado de dados armazenados em grandes bases de informações (FAYYAD e outros, 1996).

Estes autores explicam ainda que a MD é apenas parte de um processo maior de descoberta de conhecimento em bases de dados, o chamado KDD (*Knowledge Discovery in Databases*).

O processo de KDD é composto de cinco passos principais ilustrados na figura a seguir.

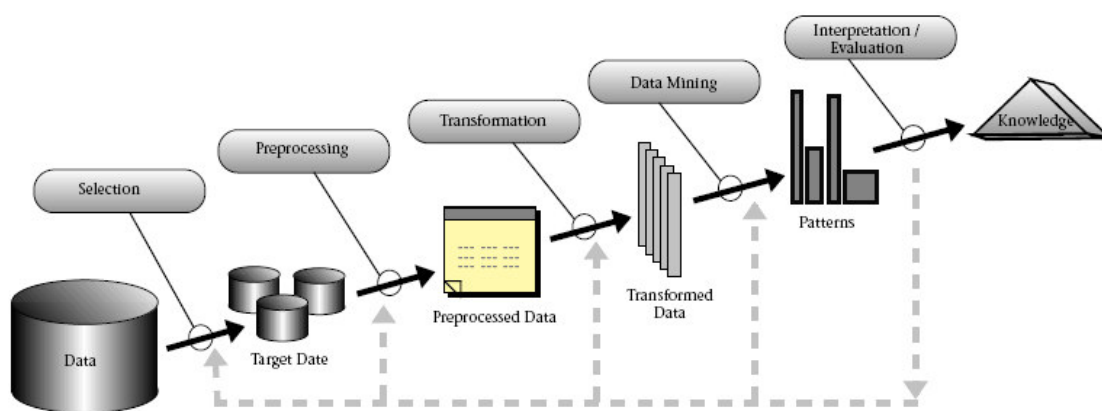


Figura 7 - Passos que compõem o processo de KDD

Fonte: Fayyad e outros (1996)

Pode-se perceber que o primeiro passo é a seleção, onde se identifica quais dados são desejados, fazendo-se uma triagem inicial, descobrindo onde eles estão armazenados, obtendo acesso aos mesmos e trazendo-os para o repositório de dados, caso necessário.

Após este passo, muitas vezes se torna necessário limpar os dados encontrados e tomar algumas decisões sobre como tratar campos que são capturados sem informações, por exemplo, esta é a etapa de pré-processamento.

O passo seguinte corresponde à transformação dos dados, para que estes possam ser utilizados pelos algoritmos de mineração de mineração, por exemplo, reduzindo o número de variáveis, encontrando representações de dados invariáveis ou convertendo medidas entre variáveis para melhorar o processo de análise posterior.

O próximo passo é propriamente aplicar um dos métodos de mineração de dados sobre os dados transformados, por exemplo, sumarização, classificação, regressão, aglomeração e demais métodos (FAYYAD; PIATETSKY-SHAPIRO; SMITH, 1996b).

O processo termina com a análise e interpretação sobre o que se obteve de resultado com a mineração dos dados, nesta fase a informação minerada é transformada finalmente em conhecimento.

Analisando este processo, pode-se perceber qual a complexidade para se descobrir conhecimento em bases de dados estruturadas. Tendo isso em vista é importante ressaltar que o surgimento da internet ampliou as possibilidades de armazenamento e de disponibilidade de informações de forma *on-line*.

4.2 MINERANDO DADOS EM FORMA DE TEXTO LIVRE

As diversas instituições do contexto social produzem e armazenam muitos documentos como relatórios, atas, históricos e vários outros tipos de informações. Em meio a este cenário é possível observar que grande parte destas informações está em formato textual livre, ou seja, escrita em linguagem natural, e que extrair informações úteis destas fontes de dados, torna-se uma tarefa ainda mais complexa do que a do KDD.

Neste sentido, as pesquisas feitas sobre MD geraram uma sub-área chamada Mineração de Textos (MT), a qual se utiliza de muitas técnicas de MD para revelar automaticamente padrões e informações úteis dentro de texto não estruturados, em linguagem natural (EBECKEN; LOPES; COSTA apud RESENDE, 2005).

Piatetsky-Shapiro e outros (2006) explicam que a Mineração de Textos é uma área de pesquisa que tenta resolver o problema de sobrecarga de informações usando técnicas de mineração de dados, aprendizado de máquina, processamento de linguagem natural (NLP), recuperação de informação (IR) e gestão do conhecimento.

Dentre essas técnicas citadas, Konchady (2006) complementa que apesar da MT ser um termo relativamente novo, ela depende muito das pesquisas feitas nas áreas de IR e NLP.

A história da IR data da década de 1960 e tem o foco em consultas do tipo booleana parecidas com as que existem atualmente em sítios de busca da web. As

pesquisas nesta área se concentram atualmente em criar bases de dados textuais e incremento de desempenho e conectividade, porém poucos sistemas de IR começaram a prover funcionalidades mínimas de busca através de linguagem natural (KONCHADY, 2006, p.4).

Portanto, um sistema de IR pode ser entendido como um sistema de busca por palavras-chave e expressões booleanas o qual exige um entendimento por parte do seu usuário sobre como montar expressões que gerem seus resultados satisfatoriamente.

A NLP, por sua vez, nasceu como uma sub-área da Inteligência Artificial (IA), com o intuito inicial de se criar uma máquina que pudesse se comunicar em linguagem natural. A idéia básica é resolver dois tipos de problemas: entender e gerar linguagem natural, porém os estudos no sentido do entendimento da linguagem natural se mostraram mais difíceis do que se previa (KONCHADY, 2006 p.5).

Em situações como as que serão estudadas no presente trabalho, onde se necessita entender o que uma pessoa quis dizer em seu texto para tomar alguma decisão posterior, as técnicas de NLP podem se configurar um caminho a ser estudado. Entretanto, Konchady (2006) alerta que esta área tem centrado seus esforços na análise ou síntese dos textos naturais e não necessariamente no entendimento profundo destes.

O autor afirma por fim que a Mineração de Textos não deve ser entendida como uma substituta da IR ou da NLP, mas como um caminho para responder as questões que ficam em aberto nestas duas abordagens.

O processo da MT se assemelha em alguns aspectos ao da MD e pode ser entendido em três etapas básicas conforme é visto a seguir na figura 8, adaptada de Ebecken, Lopes e Costa (apud RESENDE, 2005):

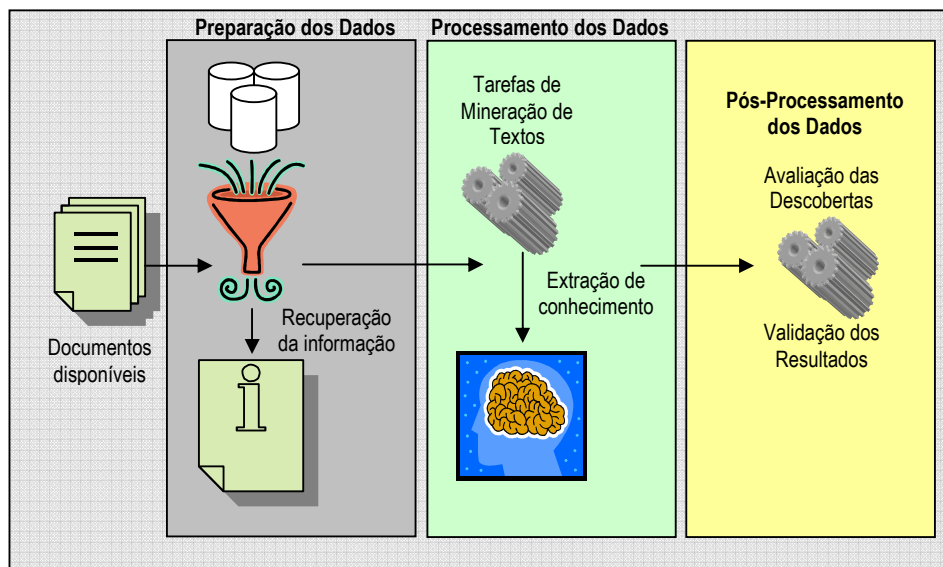


Figura 8 - Etapas da mineração de textos

Nota: Adaptado de Ebecken, Lopes e Costa (apud RESENDE, 2005).

A primeira etapa diz respeito ao pré-processamento dos dados e consiste em selecionar os dados que formarão a base de textos de interesse e o descarte das informações que não irão refletir idéias consideradas importantes. A segunda etapa é a do processamento dos dados, a qual se utilizará de tarefas de MT, tais como indexação, categorização ou sumarização para se conseguir o objetivo da extração do conhecimento sobre os textos. Por fim, tem-se a etapa de pós-processamento, a qual se concentra na validação das descobertas efetuadas pela etapa anterior e pela visualização dos resultados encontrados (EBECKEN; LOPES; COSTA apud RESENDE, 2005).

Esses autores destacam ainda a existência de dois tipos principais de abordagens aos dados em formato de texto, são eles: a análise semântica, e que se baseia na funcionalidade dos termos dentro dos textos, e a análise estatística, a qual se baseia na freqüência de ocorrência dos termos.

É importante ressaltar que estas duas abordagens podem ser utilizadas em conjunto, mas para o presente trabalho será interessante estudar a primeira delas. Esta tem relação direta com um dos objetivos da pesquisa, o qual busca o entendimento sobre o que se está falando em cada postagem dos usuários de grupos de discussões.

4.3 ANÁLISE SEMÂNTICA (ENTENDENDO UM TEXTO)

A idéia de se criar meios para que uma máquina entenda texto em linguagem natural não pode ser considerada uma tarefa simples, visto que o grau de complexidade implicado neste processo exige muita minúcia. Alguns detalhes a serem observados neste contexto, podem passar despercebidos pelos algoritmos existentes.

Um exemplo que pode esclarecer melhor esta afirmação é a análise da seguinte frase: “A manga estava bem verde.” Quando alguém tenta interpretar uma frase como esta, precisará entender o contexto no qual a pessoa está inserida. Inicialmente, pode-se interpretar esta situação por dois prismas: a palavra manga pode ser aqui entendida como sendo uma fruta ou como sendo, por exemplo, manga de uma camisa que alguém pode ter sujado de uma tinta verde.

Outro exemplo pode ser visto na pergunta: “Qual é a situação de Salvador?”. Nesse caso, se Salvador for uma cidade, a situação pode remeter uma circunstância sócio-econômica, política, climática, dentre outras similares. Caso Salvador seja entendido como um nome de uma pessoa, a pergunta pode estar se referindo ao seu estado de saúde, problemas financeiros e/ou questões emocionais, por exemplo.

Em ambos os casos mencionados, o eixo de análise está centrado no entendimento do contexto sob o qual a frase foi feita, ou seja, reconhecer que um pequeno texto está inserido em um contexto. A ocorrência destas situações de comunicação com interpretações ambíguas é associada ao conceito de polissemia, ou seja, uma palavra que possui vários significados a depender do contexto da frase em que se encontra (KONCHADY, 2006).

Assim, o desafio se traduz em fazer uma máquina conseguir entender o contexto e captar o real significado da palavra na frase.

Outro aspecto de igual relevância no entendimento semântico de um texto remete à idéia da sinonímia, ou seja, palavras diferentes que podem ser utilizadas com um mesmo significado. Ao criar um algoritmo para entendimento de um texto é importante para o seu desempenho que ele possa tratar os sinônimos como sendo apenas uma palavra, encurtando assim a quantidade de comparações necessárias para suas tarefas de mineração.

Konchady (2006) ressalta ainda que, com o passar dos tempos, alguns símbolos foram criados e são freqüentemente usados em substituição a palavras ou até mesmo a frases inteiras em um texto, são os chamados *emoticons*, os quais se constituem de caracteres para representar emoções no texto que revelem alegria, tristeza, raiva, espanto, dentre outros.

O uso de *emoticons* se revela um complicador para o ato de decodificar mensagens expostas em espaços virtuais. O advento da sociedade digital, que conduz os sujeitos a interagirem com formas de comunicações ágeis é uma demanda que se instalou e modificou formas de elaborar o pensamento (LEVY, 1996).

Esta forma de ler, interpretar e reelaborar o mundo gera uma subjetividade significativa nos espaços de interação virtual e, em especial, nos espaços de aprendizagem coletiva, que demandam a elaboração e compartilhamento de conceitos. Para algoritmos que tentam interpretar textos, esta circunstância se configura um entrave, já que a máquina possui limites consideráveis quanto à intencionalidade da subjetividade humana.

4.4 CATEGORIZANDO TEXTOS

Categorizar textos consiste em uma tarefa muito utilizada por ambientes como livrarias ou bibliotecas, onde se necessita organizar os documentos/livros em múltiplas categorias, com a finalidade de facilitar o acesso posterior a esses documentos. Para tanto, segundo Konchady (2006), tem-se um conjunto de n categorias para as quais se almeja associar uma quantidade m de documentos. Tal situação pode ser representada através da seguinte figura:

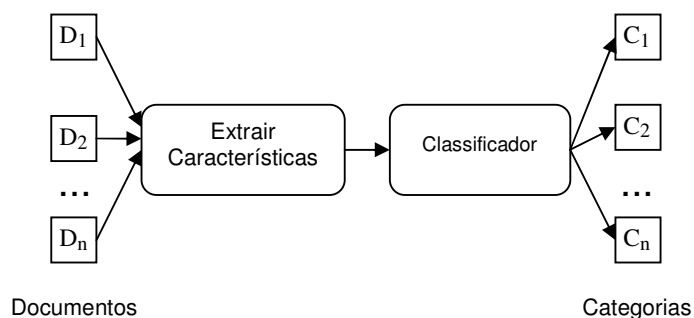


Figura 10 - Associando documentos a categorias

Nota: Adaptado de Konchady (2006, p.300).

Nesta abordagem, as categorias são diferenciadas através de palavras-chaves predefinidas associadas a elas, as quais podem ser entendidas então como uma espécie de descrição da categoria. Este autor sinaliza que “o processo de identificação dessas palavras-chaves é chamado de extração de característica” (KONCHADY, 2006, p. 300).

Dando continuidade a esta análise, Konchady chama atenção para alguns problemas/questões a serem enfrentados neste método de mineração de textos, são eles:

- a) Analisar quantas categorias seriam suficientes dada uma coleção de documentos.
- b) Qual o tamanho máximo de uma categoria (quantidade de suas palavras chaves)?
- c) Organizar as categorias de forma hierárquica será melhor do que de forma plana (tabular)?
- d) Definir se os documentos podem ser associados a mais de uma categoria ou não.

O objetivo principal da solução proposta neste estudo consiste em buscar especialistas a partir da análise de mensagens que estes postem em grupos de discussão (fóruns) e, para que esta ação seja viável, faz-se necessário entender sobre o quê a mensagem postada está falando. Portanto um dos desafios do módulo a ser criado no Moddle é a categorização automática de cada mensagem postada frente a algum assunto previamente cadastrado na ferramenta.

Os estudos sobre Mineração de Textos têm revelado algoritmos capazes de efetuar análises sobre os documentos, visando identificar uma categoria que mais se aproxime destes e explicar sobre o que se trata seu conteúdo. Uma abordagem comparativa mais aprofundada sobre alguns desses algoritmos pode ser encontrada em Yong-Feng e Yan-Ping (2004) e em Bastos (2006).

Para que tais algoritmos funcionem, Bastos (2006) destaca que é fundamental a etapa de pré-processamento, a qual possui o objetivo principal de identificar cada termo que compõe o texto. Esta fase é comum a muitos problemas da mineração de textos e, geralmente, é composta de três atividades principais destacadas pela autora, são elas:

- a) A “Tokenização” ou análise léxica, que se utiliza de um *parser* para identificar cada termo do texto como sendo uma seqüência de letras e/ou dígitos.
- b) A análise morfológica, efetuada por um *stemmer*, o qual faz uma redução de cada termo ao seu radical comum, evitando assim processamento desnecessário de palavras que possuem um mesmo valor semântico para o texto. Por exemplo: andar, andando e andado, todas elas se remetem ao ato de andar, identificado pelas letras reveladas no seu radical, portanto, poderiam ser interpretadas da mesma forma.
- c) A retirada das *stopwords*, ou seja, retirada de palavras/termos que não agregam nenhum valor/informação útil à tarefa posterior de interpretação do texto, tais como: preposições, conjunções, artigos, etc.

Após essas atividades o texto pode ser considerado pronto para a execução dos algoritmos de mineração sobre o seu conteúdo. Entretanto, os horizontes da presente pesquisa se delimitam no enquadre específico da análise de situações referentes à categorização automática de textos.

Um ponto interessante a ser destacado é o fato de que os textos analisados ao longo desta pesquisa são mensagens de curto comprimento e a maioria dos algoritmos pesquisados se concentra em grandes documentos e volumes de dados.

Portanto, decorrente da análise feita neste estudo, será oferecida a comunidade acadêmica uma possível alternativa à categorização das postagens dos fóruns, levando-se em consideração as peculiaridades supracitadas, bem como as características, dos algoritmos já existentes, que mais se adéquam ao processamento de mensagens deste tamanho.

Wang e Taylor (2007) destacam dois métodos que são frequentemente usados para categorização de textos, sendo eles o *Vector Space Model* – VSM (Modelo do Espaço Vetorial) e o *Latent Semantic Indexing* – LSI (Indexação Semântica Latente), cada qual com suas vantagens e desvantagens. Um aspecto importante a ser observado nestes dois métodos é que eles são baseados em buscas de documentos por palavras-chave e, portanto, convergentes com os interesses desta pesquisa.

O presente trabalho faz uma exploração maior sobre o VSM, visto que a LSI se utiliza de algoritmos e técnicas mais rebuscadas, porém com utilidade em volumes maiores de textos do que os de mensagens postadas em fóruns, já que se propõe a procurar documentos relevantes sem precisar examinar documentos inteiros (RESENDE, 2005).

Segundo Magalhães (2008), no Modelo do Espaço Vetorial, cada documento pode ser visto como um vetor de termos e cada um destes possuem um peso associado que indica o seu grau de importância dentro do documento. Tal modelo poderia ser representado da seguinte forma:

Vetor do Documento: $d_j = \{w_{1j}, \dots, w_{|t|j}\}$, onde:

d_j é a representação de um documento e w_{ij} representa o peso associado a um dado termo de um conjunto de t termos do referido documento.

Quadro 2 - Representação de um Vetor de Termos de um Documento

Fonte: Magalhães (2008).

Tendo em vista a elucidação da representação acima, segue uma situação exemplo:

Trecho de texto (extraído de Resende, 2005, p.346):

“Apesar de sua simplicidade, o Modelo do Espaço Vetorial (VSM) e suas variantes são freqüentemente a forma mais comum de representar documentos textuais na mineração de coleções de documentos...”

Quadro 3 - Exemplo de representação de um texto em relação aos seus termos

Alguns termos extraídos e seus respectivos pesos (entre parêntesis) definidos pela quantidade de ocorrência destes dentro do texto:

simplicidade (1), variantes (1), documentos (2), mineração (1), coleções(1)

Quadro 4 - Exemplo de representação de um texto em relação aos seus termos

Observa-se que a palavra “documentos” ganhou um peso maior dentro do documento, pois esta aparece mais vezes do que as demais.

Segundo Salton (1999 *apud* RESENDE, 2005, p. 346), no Modelo do Espaço Vetorial “os documentos são representados como pontos (ou vetores) em um espaço Euclidiano t -dimensional em que cada dimensão corresponde a uma palavra (termo) do vocabulário”. A figura 11 traz uma ilustração desta estratégia de representação proposta por Salton (1975).

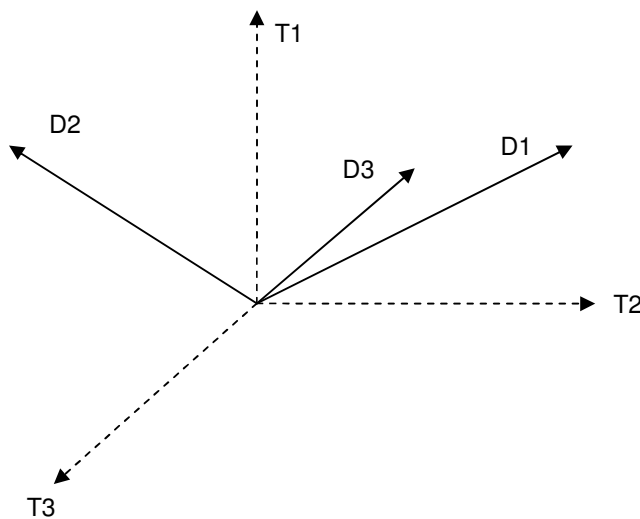


Figura 11 - Modelo do Espaço Vetorial.

Nota: Adaptado de Salton (1975).

Pode-se observar que o documento D1 está mais próximo do D3 em relação ao documento D2, pois o ângulo que os separa é visivelmente menor. Um algoritmo que se utilize os conceitos do VSM, se comporta desta forma, pois, segundo Resende (2005), a idéia principal é que os documentos sejam considerados tão semelhantes quanto menor for o ângulo de diferença entre eles no espaço

Euclidiano t -dimensional (representado acima), desconsiderando o tamanho do documento (RESENDE, 2005).

No caso do presente trabalho, a idéia do VSM se mostrou mais adequada frente à do LSI e pôde ser adaptada da seguinte maneira:

- a) As palavras chaves são os termos que compõem a mensagem postada (extraídos pelo algoritmo de tokenização).
- b) Os documentos são as categorias, cada qual composta pelos termos que as descrevem.

O intuito final é identificar se a coleção de termos de uma mensagem postada (ou seja, o seu vetor de termos), se aproxima de alguma categoria específica (ou do vetor de termos os quais a compõem).

Os autores destacam, todavia, que uma das desvantagens do uso do VSM é a inabilidade de tratar os problemas de polissemia e sinonímia, conceitos discutidos anteriormente.

Para o problema da sinonímia, o módulo aqui proposto, intitulado como **FindYourHelp**, utiliza um algoritmo para confrontar os termos da mensagem e das categorias com uma base de dados de sinônimos, esta estrutura é chamada comumente de *thesaurus*, e pode ser definida como “um dicionário que lista palavras as quais possuem um significado similar ou relacionado.” (NABER, 2004, p.54).

O problema de polissemia, por sua vez, não será abordado pelo módulo devido a uma decisão de projeto, a fim de preservar a performance do funcionamento do ambiente Moodle, no momento em que cada postagem é efetuada.

Alguns autores recomendam ainda utilizar a técnica do VSM juntamente com algum mecanismo de ponderação dos termos (*weighting*), de forma a evitar algumas situações onde as mensagens podem ser categorizadas inadequadamente (de fato, o próprio modelo vetorial já pressupõe o uso de pesos para seus termos conforme citado anteriormente neste tópico).

Tal situação é verificada, por exemplo, quando uma categoria é definida por um conjunto menor de termos de que outra categoria (*vide* demonstração a seguir).

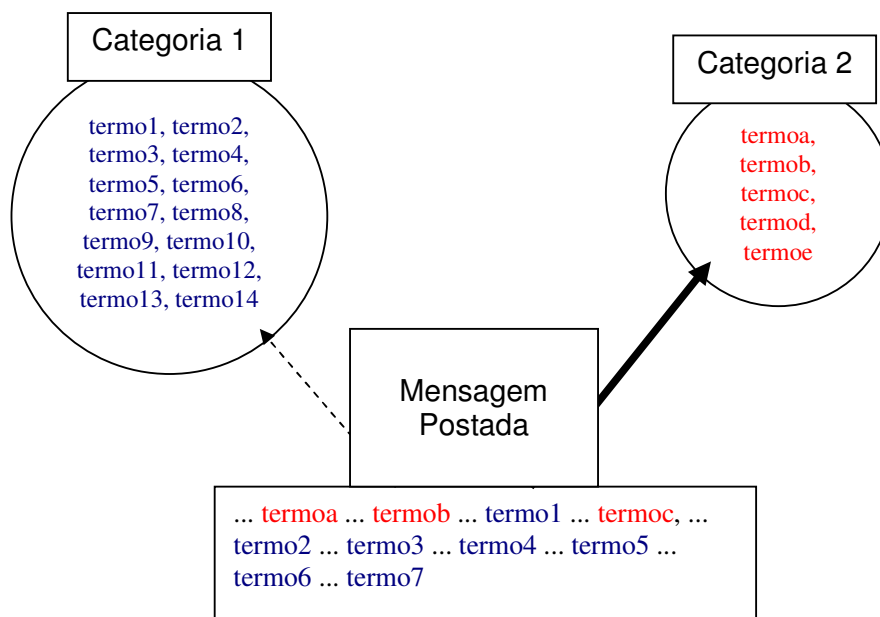


Figura 12 - Categorização de mensagens sem a utilização de um método de ponderação

Neste caso, pode-se perceber que a mensagem postada fala mais sobre a categoria de maior quantidade de termos (categoria 1), mas também se refere a termos da categoria que possui um conjunto menor de termos (categoria 2). Entretanto, através de um cálculo simplista, onde se leva em consideração apenas a quantidade de termos da mensagem sobre o total de termos da categoria, conclui-se que a mensagem poderá ser categorizada erroneamente, como pertencente à categoria de quantidade menor de termos (*vide* as setas da figura), por exemplo:

- a) 1º Cálculo → A mensagem postada possui 7 termos pertencentes ao grupo de 14 termos da categoria 1. Portanto, dividindo-se os 7 termos pelos 14 da categoria obtém-se um valor igual a 0,5, que neste caso pode ser entendido como sendo uma aproximação de 50% da mensagem para com esta categoria.
- b) 2º Cálculo → A mensagem postada possui 3 termos pertencentes ao grupo de 5 termos da categoria 2. Assim, dividindo-se os 3 termos pelos 5 da categoria obtém-se um valor igual a 0,6, que neste caso pode ser entendido como sendo uma aproximação de 60% da mensagem para com esta categoria.

- c) Conclui-se: um algoritmo que leve em consideração apenas tais informações, decidiria que a mensagem está mais próxima da categoria 2, embora, visualmente, é perceptível que a mensagem tem maiores chances de estar abordando assuntos relacionados à categoria 1.

A proposta do algoritmo do módulo **FindYourHelp**, discutida com maiores detalhes no capítulo posterior, vem contornar tal problema aplicando um método para ponderação dos termos encontrados na mensagem, também chamado de TFIDF (*Term Frequency Invert Document Frequency*).

Segundo Soucy e Mineau (2005), este método de ponderação é o mais comum para o modelo espaço vetorial, particularmente utilizado em contextos com problemas que envolvem a recuperação de informações.

A abordagem TFIDF é argumentada por Robertson (2004), onde explica que inicialmente uma função de ponderação de termos chamada *inverse document frequency*, mais conhecida como IDF, foi proposta por Karen Spark Jones (1972 *apud* ROBERTSON, 2004). Essa função se baseia na contagem de documentos presentes em uma coleção, os quais possuem um determinado termo consultado (ROBERTSON, 2004). Face ao exposto, a situação se apresenta da seguinte forma: se um termo consultado ocorre em muitos documentos, este termo não é considerado suficientemente bom para poder diferenciá-los, portanto um peso menor deve ser atribuído a este termo.

O autor complementa que essa teoria representou um grande salto no campo da recuperação de informação. Destaca ainda que a IDF quando associada à medida do *Term Frequency* (TF), ou seja, a quantidade de vezes que um termo aparece dentro do mesmo documento, gerou uma função de ponderação de termos extraordinariamente robusta e difícil de ser superada por outras teorias, nela multiplica-se a medida TF pela medida IDF. Tal abordagem é utilizada por quase todos os esquemas de ponderação de termos e é conhecida como TF*IDF, justamente por conta multiplicação aplicada em suas medidas.

A fórmula básica para cálculo do IDF pode ser entendida da seguinte forma: assumindo-se que existam N documentos e que o termo t_i ocorre em n_i deles, então,

a medida proposta por Sparck Jones (apud ROBERTSON 2004) pode ser representada como se segue:

$$idf(t_i) = \log (N/n_i)$$

Logo, multiplicando-se a frequência do termo dentro de um mesmo documento pela medida IDF referente a esse termo, tem-se a medida TF*IDF deste termo, ou seja, o seu peso em relação ao documento. Conforme já dito anteriormente, esta função de ponderação será utilizada pelo algoritmo do módulo **FindYourHelp**, como forma de aprimorar a técnica do VSM.

4.5 CONSIDERAÇÕES FINAIS

Muitos dados estão contidos em grandes bases de informações dentro das instituições, mas considera-se difícil extrair conhecimento destas bases apenas com recursos da capacidade humana. A mineração de dados é uma vasta área de conhecimento e muitas pesquisas são feitas sobre este tema de forma a aperfeiçoar tal tarefa.

Dentre essas pesquisas, nota-se que as formas de se encontrar informações úteis em dados armazenados em formato de texto livre possuem características específicas, as quais merecem uma atenção diferenciada, o que acabou por criar uma subárea dentro da mineração de dados: a Mineração de Textos.

Esta subárea foi analisada neste capítulo com o intuito de ilustrar algumas possibilidades de tratamento das mensagens veiculadas em fóruns de discussão, a fim de se encontrar especialistas nestes espaços. Dentre as diversas técnicas estudadas, pode-se destacar a categorização de textos, cuja funcionalidade servirá de ponto de partida para a solução proposta neste trabalho.

O próximo capítulo descreve o módulo **FindYourHelp**, que foi desenvolvido tendo como base a proposta de busca de especialistas em fóruns do AVA Moodle apresentada no capítulo sobre a ferramenta. Tal solução utilizará algoritmos de mineração de textos, que abordam, mais especificamente, o processo de categorização de textos.

5 FINDYOURHELP: BUSCANDO ESPECIALISTAS DENTRO DE UM AVA

Dentre as principais territorialidades de construção de saber atuais estão os Ambientes Virtuais de Aprendizagem. Tais ambientes vinculam pessoas e entrelaçam saberes a partir de criação de tópicos de discussão, postagens de mensagens em fóruns, chats on-line, criação coletiva de documentos (WIKI), dentre outras funcionalidades.

Um dos grandes diferenciais da aprendizagem em AVA pode ser entendido a partir do pressuposto de que a aprendizagem ganha um significado diferenciado quando se apóia na ação colaborativa, pois reflete pessoas que mutuamente se ajudam voltadas para a construção e reconstrução de saberes (MOORE; KEARSLEY, 2008).

Nonaka e Takeuchi (1997) lembram que a boa gestão do conhecimento depende diretamente do tratamento que é dado às informações que circulam dentro das instituições, mas que o cerne da questão está de fato nas pessoas que produzem conhecimento e ajudam (ou atrapalham) na disseminação desta produção.

Diante deste panorama, percebe-se a importância de se buscar formas mais eficientes de gerenciar pessoas e informações inseridas em um AVA, para tirar um melhor proveito da colaboração inerente deste tipo de ferramenta. No capítulo que discute ambientes virtuais de aprendizagem, verificou-se quais desses promovem formas de se buscar especialistas atuantes em fóruns de discussão, cadastrados nesses ambientes.

Percebe-se então, após uma análise, que dentre esses ambientes virtuais muitos possuem formas de manter um perfil dos participantes, o que poderia servir de um ponto de partida inicial para a busca de especialistas dentro de um grupo. Um diferencial, desta pesquisa complementa esta idéia da simples verificação de um perfil, propondo uma busca, onde os possíveis especialistas são identificados e pontuados a partir de análises feitas sobre a sua participação, ou seja, sobre sua contribuição com o grupo em forma de postagens em fóruns de discussão.

O presente capítulo apresenta, portanto, o **FindYourHelp**, um módulo construído à luz da solução proposta pelo autor deste trabalho. Em outras palavras, uma forma de analisar as postagens dos fóruns de pesquisa dentro de um AVA,

para que estas sejam categorizadas, utilizando-se de algoritmos de mineração de textos, a fim de pontuar os autores das postagens conforme estas sejam inseridas no ambiente.

O AVA escolhido para a confecção do FindYourHelp foi o Moodle e alguns fatores motivaram a escolha por este ambiente virtual de aprendizagem, tais como: sua ampla utilização em instituições de ensino espalhadas pelo mundo; o fato de ser um software livre, o que possibilita uma liberdade de uso, distribuição e modificação do seu código fonte por parte de terceiros; e sua facilidade de extensão percebida pela grande quantidade de módulos (plugins), que já foram desenvolvidos para o ambiente (MOODLE, 2008c).

Entende-se como principais usuários deste módulo proposto, os professores e coordenadores de cursos e grupos de pesquisas que utilizam o Moodle como ferramenta de apoio aos seus projetos e que, por ventura, possuam o desejo de observar as pessoas mais participativas, bem como saber em quê elas mais contribuem dentro de cada curso/grupo. Em outro nível, o próprio participante do grupo que precisar de auxílios específicos, pode se valer deste módulo para encontrar de forma mais rápida um “especialista” no assunto que desejar.

Tendo o intuito de promover um melhor entendimento sobre o **FindYourHelp** este capítulo se destina ao detalhamento do módulo proposto onde, inicialmente, serão destacados os seus principais requisitos funcionais. Em seguida, a sua especificação técnica será abordada e, por fim, serão detalhadas a sua arquitetura, suas funcionalidades e formas de utilização por parte dos usuários.

5.1 REQUISITOS FUNCIONAIS

O módulo **FindYourHelp** foi pensado com a intenção de suprir os principais requisitos funcionais a seguir:

- a) Exibir a hierarquia de categorias definidas previamente em uma tabela à parte e permitir sua atualização por parte dos professores/coordenadores de curso/grupos de discussão;
- b) Analisar as postagens das mensagens em fóruns assim que estas forem enviadas pelos usuários, categorizando-as dentro da hierarquia

definida (requisito 1). Esta função do módulo não possui interação direta com o usuário, o **FindYourHelp** aproveita o momento em que a mensagem é postada no fórum, para capturá-la internamente e proceder com a análise do texto;

- c) Identificar e buscar especialistas em determinados pontos da hierarquia de categorias, permitindo o acesso aos perfis correspondentes.

Tais requisitos podem ser decompostos em atividades menores, as quais serão aqui descritas de modo a evidenciar o processamento feito pelo módulo em cada situação. Serão destacados também os pontos em que são aplicados passos necessários para executar a mineração de textos, conforme se segue:

- a) Atividades do requisito 1:
 - o Exibir categorias: esta é a tela inicial deste requisito, a qual exibe as categorias em formato de árvore, tal qual a estrutura comumente utilizada para exibição de diretórios em sistemas operacionais comuns de mercado, como a família do Microsoft Windows ou a família do Linux.
 - o Adicionar novas categorias/termos: o professor/coordenador de um grupo de estudantes pode criar categorias hierarquicamente dependentes, informando para isto seu nome e uma descrição opcional. Conforme já mencionado anteriormente, a organização dos itens se dá tal como uma árvore, onde os elementos mais específicos (folhas ou termos) descrevem os elementos mais genéricos (raízes ou categorias).
 - o Editar/remover Categorias/Termos: ao clicar em uma categoria/termo o usuário pode optar por editar seu nome/descrição ou por excluí-la, sendo assim uma tela de edição será exibida no primeiro caso e uma tela de confirmação de exclusão será exibida no segundo caso.

b) Atividades do requisito 2:

- Captura e Remoção de Pontuação: O algoritmo captura a mensagem no momento em que esta é postada no fórum e procede com a remoção dos sinais de pontuação e de números arábicos (0, 1, 2...), porém levando em consideração palavras separadas por hífen, estas se mantêm intactas. Uma variável do tipo *string* é gerada agora com o texto sem os sinais de pontuação e sem os números.
- Tokenização: após a atividade anterior, o algoritmo transforma o texto em um vetor de termos, utilizando o caractere que representa um espaço em branco como delimitador das palavras/termos.
- Remoção de *stopwords*: conforme já descrito anteriormente, algumas palavras (*stopwords*) são removidas do vetor de termos por não fazerem diferença alguma na análise da mensagem. Para tanto, foi utilizado um conjunto de palavras freqüentes em português retiradas do sítio da Linguateca (LINGUATECA, 2009).
- Execução dos processos de *Stemming* e redução de sinônimos: esta parte do algoritmo cuida de fazer o seguinte processamento: para cada palavra encontrada no vetor de termos, aplicar a redução ao seu radical comum (*stem*) e, em seguida, aplicar a redução ao sinônimo comum utilizando-se de um thesaurus em português extraído do projeto OpenThesaurusPT (OPENTHESAURUS, 2009). O resultado deste processamento é um termo reduzido o qual será posteriormente utilizado na comparação com os termos das categorias
- Categorização da Mensagem: esta é a parte que se concentra em criar um vetor de similaridades, entre a mensagem e as categorias existentes, utilizando-se da técnica aplicada no Modelo do Espaço

Vetorial, discutido anteriormente, aplicando o cálculo do TF/IDF como forma de ponderação dos termos

- Pontuação do participante no *ranking* de especialistas: aqui o algoritmo faz uma análise para verificar qual a categoria do vetor de similaridades mais se aproxima da mensagem postada e, por conseguinte, procede com a pontuação do participante no ranking de categorias, atribuindo um ponto para este participante associado à categoria escolhida
- c) Atividades do requisito 3:
- Validação de especialistas: uma listagem prévia dos especialistas encontrados pela ferramenta pode ser validada por participantes com perfis de professores ou coordenadores cadastrados no Moodle, a fim de promover uma validação das descobertas do algoritmo de categorização em uma forma humanamente assistida.
 - Listagem de especialistas: a árvore de categorias também dá a oportunidade de buscar os participantes considerados especialistas em uma dada categoria. Para isto foi criado um esquema de grupos (A, B e C) buscando representar quem mais contribuiu na ferramenta do fórum, até o momento, com postagens sobre a categoria escolhida (grupo A), quem contribuiu de forma moderada (grupo B) e quem contribuiu menos até o momento (grupo C). Cabe aqui uma observação, a qual será detalhada mais adiante, sobre as duas formas de exibição desta listagem: a primeira é uma listagem simples que indica igualdade entre os participantes que nela aparecem, ela sempre será exibida caso o usuário conectado no Moodle seja um estudante; a segunda se trata de um caso especial, onde o usuário corrente é do perfil Professor, nesta situação haverá uma diferenciação visual para o caso dos participantes que foram validados na atividade explicada no item anterior.

- Exibição do perfil do especialista: a listagem de especialistas traz consigo um *link* apontando para o perfil de cada participante (funcionalidade que já existe no AVA Moodle e foi aqui aproveitada pelo FindYourHelp), a fim de permitir o contato com a pessoa de modo facilitado.

Observando os itens acima, evidencia-se que as atividades ligadas à mineração de textos do módulo **FindYourHelp** se concentram naquelas ações que compõem a análise automática da postagem no fórum.

Para aprofundar esta análise serão apresentadas e explicadas algumas características técnicas sobre o módulo **FindYourHelp**.

5.2 CARACTERÍSTICAS TÉCNICAS

Um dos principais elementos que caracterizam o presente trabalho se traduz no fato do **FindYourHelp** ser inicialmente pensado como uma implementação de um módulo do AVA Moodle e, neste caso, algumas regras foram necessariamente impostas ao desenvolvimento do mesmo, tais regras podem ser entendidas como requisitos não-funcionais de implementação:

- a) A linguagem PHP: O ambiente Moodle foi construído utilizando a linguagem PHP, e, portanto, para que se possa criar um módulo aderente a este ambiente, o mesmo precisa ser implementado utilizando-se esta linguagem. A versão utilizada foi a 5.0, a qual possui suporte à programação orientada a objetos. Vale ressaltar que o ambiente Moodle é executado a partir de um servidor Web, que no caso do presente trabalho este foi o Apache na sua versão 2.2.
- b) O uso de recursos de interface *web* simples: o projeto do ambiente Moodle preza pelo desempenho da aplicação e pela execução de suas funções de forma que a maioria dos navegadores de internet disponíveis no mercado possa utilizar seus recursos em sua totalidade (MOODLE, 2008a). Este fato traz algumas recomendações aos

desenvolvedores para evitar o uso de recursos mais avançados, tais como: os *scripts* do lado cliente mais complexos (*javascripts*), utilizar *Web Services* em demasia, ou ainda recursos como objetos Flash (ADOBE, 2009) ou Applets Java (ORACLE, 2009) para interação com o usuário. Pela filosofia do projeto Moodle, cada página deve utilizar a biblioteca de funções do próprio ambiente, as quais tratam as requisições geralmente na forma mais entendida pelos navegadores existentes, ou seja, através de requisições ao servidor *web* e respostas deste mesmo servido em formato HTML “puro”.

- c) O banco de dados MySQL: em analogia ao item anterior, foi necessário implementar as tabelas do módulo **FindYourHelp** em seu ambiente de testes com o sistema gerenciador de banco de dados MySQL (MYSQL, 2009), para que não houvesse diferenças frente ao ambiente onde foi aplicado o estudo de caso do presente trabalho.

Tendo o objetivo de proporcionar uma melhor visualização dos processos, as funcionalidades do **FindYourHelp** serão representadas a seguir em forma de telas (quando necessárias), diagramas de atividades e diagramas de seqüência.

5.3 ARQUITETURA DO MÓDULO FINDYOURHELP

Conforme já mencionado anteriormente, o módulo proposto se integra ao AVA Moodle, portanto, entender a sua arquitetura só se torna possível quanto identificado o contexto em que este se integra ao ambiente supracitado. Uma ressalva, entretanto, se faz necessária neste momento, a fim de explicar que a arquitetura aqui pensada possibilita que a idéia central do módulo (a categorização automática das mensagens do Fórum e a classificação de especialistas) seja portada para outros ambientes virtuais de aprendizagem.

Outra explicação que se faz necessária neste momento é a escolha pela categorização de mensagens provenientes da funcionalidade do Fórum em detrimento daquelas provenientes de um *Chat*. Tal opção se deu pelo fato de que as mensagens veiculadas em um Fórum tendem a possuir uma maior qualidade em seu

texto frente às de um Chat, pois pela sua característica assíncrona, o fórum permite que os participantes tenham um tempo maior de reflexão antes de cada postagem. Tal situação, porém, não é verificada em um Chat, cuja proposta é ser um meio de comunicação síncrono (MOORE; KEARSLEY, 2008).

A figura 13 é um diagrama de componentes o qual ilustra a arquitetura atual do **FindYourHelp**, bem como sua forma de interação com o Moodle.

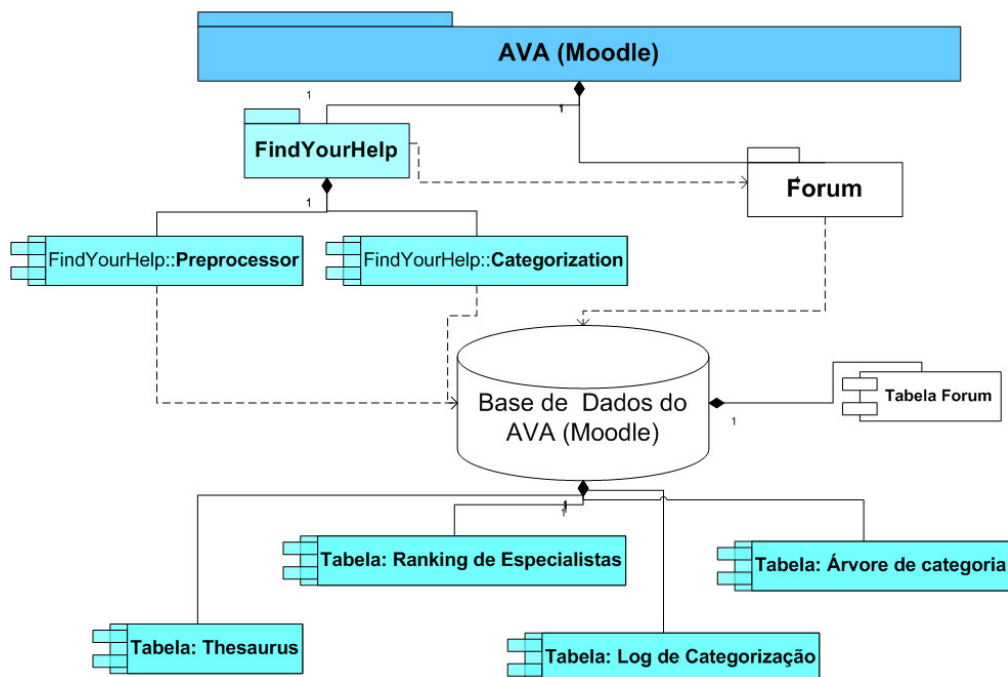


Figura 13 - Arquitetura do FindYourHelp integrada ao AVA Moodle

Primeiramente, pode-se observar na figura 13, os componentes do módulo **FindYourHelp** destacados na cor azul claro, indicando que estes fazem parte do ambiente Moodle como um módulo adicional. O **FindYourHelp** é composto por dois componentes principais o Pré-processador e o Categorização e interage com o módulo *Forum* já existente no Moodle.

Outro aspecto a ser observado na figura 13 é que a base de informações do Moodle será incrementada com algumas tabelas do módulo **FindYourHelp** (igualmente destacadas em azul claro) e uma peça indispensável neste caso é a tabela de Árvore de categorias, pois conforme será discutido mais adiante, é esta informação que torna possível o processo de categorização da mensagem e conseqüente categorização dos participantes especialistas. Além desta tabela de categorias, o módulo utiliza tabelas de sinônimos (thesaurus) e uma tabela de log de

categorização que serve como auxílio na contagem das palavras utilizadas por mensagens e análises posteriores sobre as mensagens categorizadas.

Por fim, percebe-se pela figura 13 que um *ranking* de especialistas será mantido também na forma de uma tabela adicional na base de dados do Moodle e será alimentado pelo **FindYourHelp**, a fim de proporcionar a exibição, a qualquer momento, da listagem de participantes identificados como especialistas em um determinado assunto. Tal funcionalidade também será detalhada no próximo tópico.

5.4 PRINCIPAIS FUNCIONALIDADES

Conforme já mencionado anteriormente, o **FindYourHelp** foi pensado de forma a ter uma parte das suas funcionalidades que não interage com o usuário do ambiente Moodle e uma outra que necessitará de algum tipo de interação.

O tópico posterior visa discutir as funcionalidades apresentadas, detalhando-as em forma de diagramas de atividades.

5.4.1 Exibir e atualizar a hierarquia de categorias

Esta atividade é essencial para que a categorização das mensagens postadas no fórum possa ser executada com sucesso pelo algoritmo do módulo **FindYourHelp**. Para tanto, um participante com credenciais de professor ou coordenador de curso deve selecionar a opção Atualizar Categorias/Termos do menu principal do módulo. É exibida então uma árvore de categorias existentes e seus respectivos termos ou subcategorias que as compõem. A imagem que segue traz um resumo das sub-atividades que são executadas ao inserir uma nova categoria ou termo.

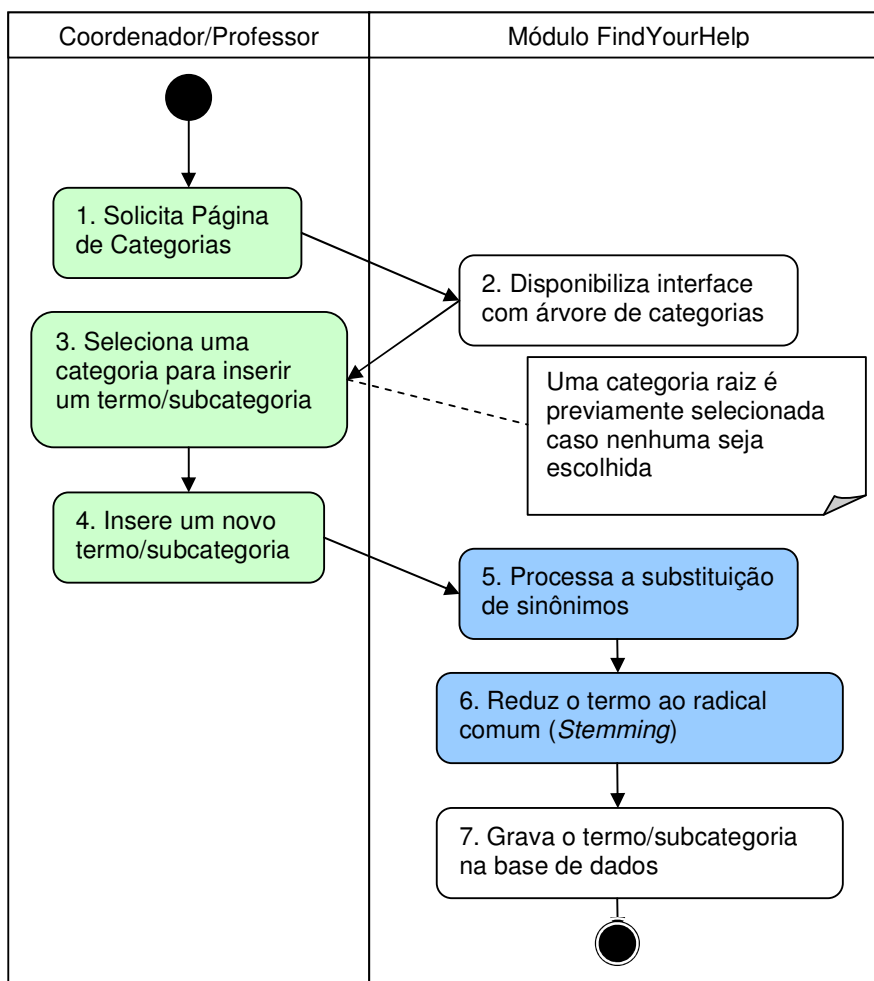


Figura 14 – Diagrama de Atividades: Inserir novo termo/categoria

O processo para inserção de uma nova categoria ou termo à hierarquia de categorias existentes começa com a solicitação da página de categorias (atividade 1), o **FindYourHelp** exibe então uma árvore de categorias e traz a opção para que o professor/coordenador possa inserir, excluir ou editar uma categoria/termo (atividade 2).

Deste ponto em diante, o usuário do sistema poderá optar pela inserção de um novo item na árvore, selecionando opcionalmente uma categoria para que esta seja definida como “pai” do termo o qual deseja inserir (atividade 3). Caso nenhuma seja selecionada, a categoria raiz é passada como parâmetro.

A continuação do processo se dá com a inserção dos dados do item e sua posterior confirmação (atividade 4), neste momento uma parte do algoritmo de pré-processamento do texto entra em ação para buscar no *thesaurus* um sinônimo para esta palavra (atividade 5). Em seguida o algoritmo processa a redução ao radical comum da palavra (atividade 6), pois este será utilizado no momento da comparação futura com as mensagens (explicado no item a seguir).

É necessário ressaltar que o algoritmo base aqui utilizado foi o RSLP – Removedor de Sufixos da Língua Portuguesa, adaptado de Coelho, Orengo e Buriol (2007) para a linguagem PHP. Por fim, o algoritmo procede com a gravação do novo termo na base de dados do Moodle (atividade 7).

5.4.2 Analisar as postagens dos fóruns

A análise das postagens é uma atividade que não possui interação direta com participantes do ambiente Moodle. Nela o algoritmo do módulo **FindYourHelp** entra em ação assim que uma mensagem for postada em um dos fóruns ativos do ambiente. Este mecanismo é detalhado na figura 15.

O processo de análise das postagens tem início com a postagem de uma mensagem em um dos fóruns ativos no ambiente Moodle (atividade 1). Neste momento o **FindYourHelp** entra em execução através de um ponto de entrada criado para proceder com a categorização da mensagem, logo após a sua inserção na base de dados do Moodle (atividade 2), ou seja, o algoritmo espera que a inserção da mensagem seja efetuada com sucesso na base para que só depois proceda com a sua análise.

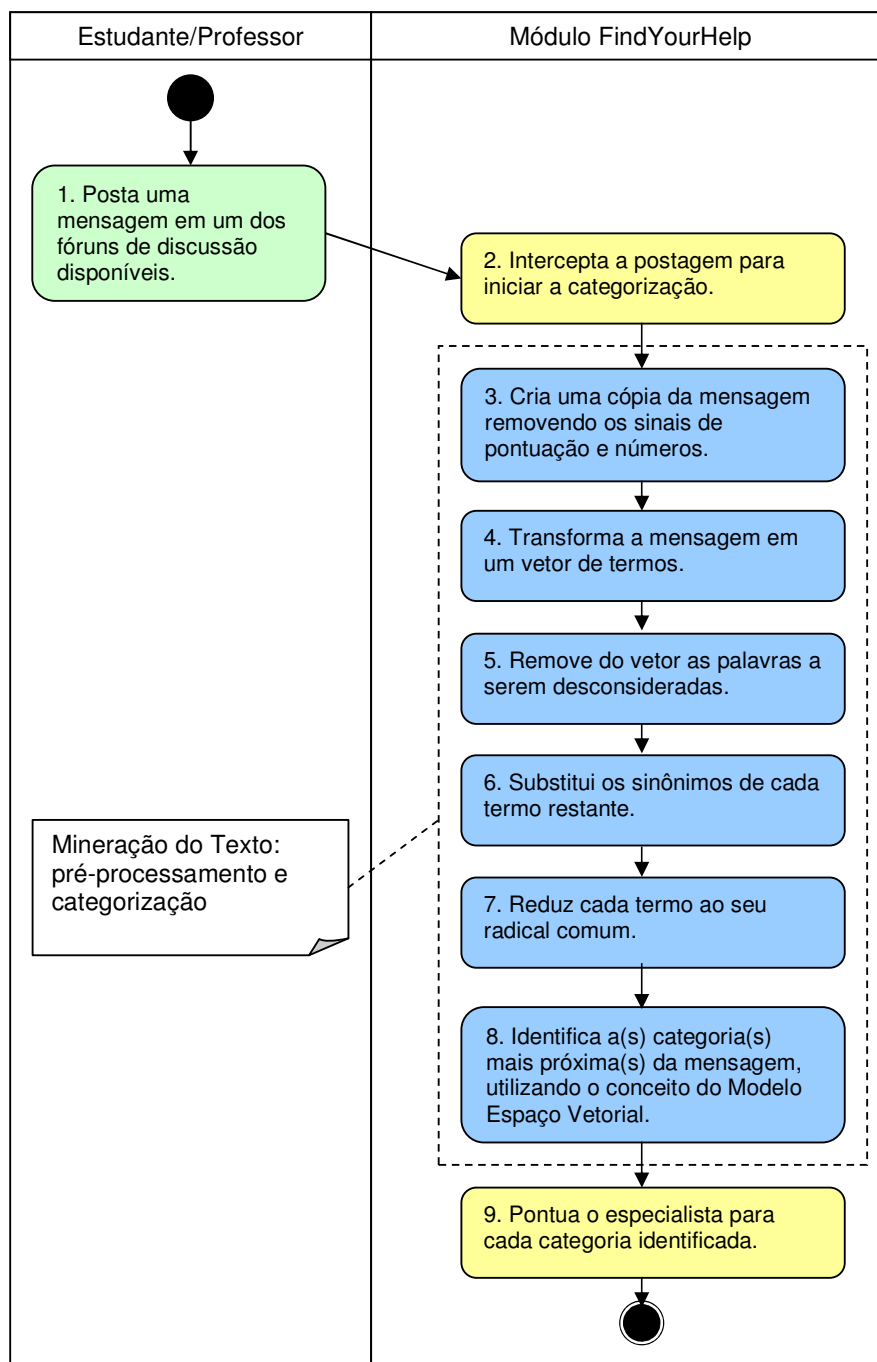


Figura 15 - Diagrama de Atividades: Análise de uma postagem no fórum

Um detalhe a ser observado é que este processo não ocorre na criação de uma nova discussão em um fórum, pois geralmente quem está criando uma discussão está fazendo uma indagação ao grupo em busca de uma resposta,

portanto, por uma decisão de projeto, o módulo **FindYourHelp** não entra em ação para proceder com a análise e categorização de mensagens desta natureza.

O próximo passo é criar uma cópia (em formato *string*) da mensagem, para que se possa remover os sinais de pontuação, tais como: vírgula, ponto, exclamação, ponto e vírgula, dentre outros e também os algarismos numéricos (atividade 3). Tais caracteres não farão diferença no momento de analisar a postagem.

Em seguida o algoritmo executa a transformação do texto em um vetor de termos, usando, para isso, o espaço em branco como caractere delimitador entre as palavras (atividade 4). Uma observação: o algoritmo, no momento da remoção de pontuações, remove os hífens que possuem espaço em branco antes e/ou depois da palavra, porém não remove os hífens que concatenam palavras compostas, desta forma, uma palavra composta será tratada como um único termo dentro do vetor.

Após a transformação da mensagem em um vetor de termos, ocorre o momento de retirar as palavras desnecessárias, ou *stopwords* (BASTOS, 2006), que são aquelas que não agregarão informação útil à análise final do texto (atividade 5). Para isto, utilizou-se uma base de cerca de cem palavras freqüentes da língua portuguesa retiradas do sítio Linguateca (LINGUATECA, 2009). Para efetuar este passo de extração de termos de um vetor foi necessário transformar as palavras do sítio Linguateca (citadas anteriormente) em itens de um vetor auxiliar e, em seguida, a remoção de cada termo presente no vetor principal que fosse equivalente ao do auxiliar.

O algoritmo segue com a substituição do sinônimo (atividade 6) para cada termo que tenha permanecido dentro do vetor de termos da mensagem após a atividade anterior. Para tanto foi necessário importar previamente, para a base de dados do Moodle, um *thesaurus* contendo as principais palavras em português e seus respectivos sinônimos. Conforme já mencionado anteriormente, estes dados foram encontrados no sítio OpenThesaurusPT (OPENTHESAURUS, 2009), como estavam em formato texto, um *software* foi desenvolvido em Java para a executar a importação adequadamente para as tabelas no Moodle (vide o código no APÊNDICE B).

É necessário ressaltar também que esta base de sinônimos pode ser incrementada e, posteriormente, adaptada para suportar termos escritos em outras línguas, visto que a estrutura conta com uma tabela contendo palavras bases e uma

relação entre os itens desta, com itens de outra tabela contendo os seus respectivos sinônimos (*vide* abaixo uma representação do modelo das tabelas).

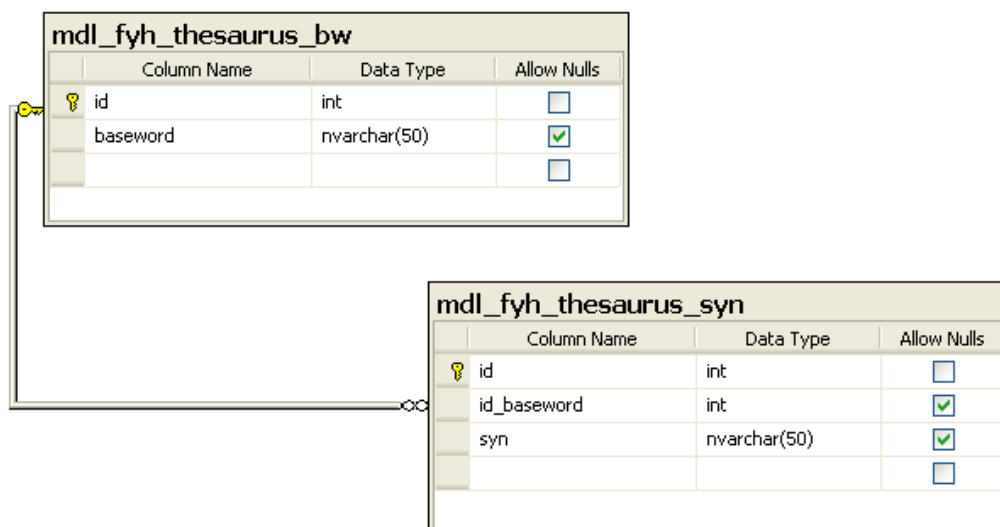


Figura 16 - Estrutura de tabelas de sinônimos do FindYourHelp

A figura 16 evidencia um relacionamento do tipo “um para muitos” entre as duas tabelas e, através de uma consulta simples no banco, é possível buscar todos os sinônimos de uma determinada palavra base cadastrada (*baseword*).

A sétima atividade do algoritmo é a redução de cada termo ao seu radical comum (processo de *stemming* – (BASTOS, 2006)). Neste caso, conforme já visto anteriormente, foi necessária uma adaptação do o algoritmo RSLP de Coelho, Orengo e Buriol (2007) para a linguagem PHP. No entanto, o algoritmo criado para este estudo foi elaborado utilizando-se um padrão de projeto chamado *Strategy*, o qual permite a mudança do comportamento de um algoritmo através da passagem de um objeto como parâmetro (GAMMA e outros, 2001). Nesta situação tal objeto deverá conter a função de “*stemming*”, logo, caso seja de interesse a troca do algoritmo, será possível realizá-la com um esforço relativamente pequeno de programação e sem maiores impactos para a página chamadora da função.

A penúltima atividade (8) é a identificação da(s) categoria(s), cujos termos mais se aproximam do conteúdo da mensagem (*bag of words*). Aqui o algoritmo implementado utiliza-se do conceito do Modelo Espaço Vetorial (SALTON, 1975), já explicado no capítulo sobre Mineração de Textos.

Para tanto, o vetor de termos da mensagem resultante da atividade 7 é tratado como o documento a ser analisado e, os termos que compõem cada

categoria, são tratados como os outros documentos do espaço vetorial a serem confrontados. Para cada documento-categoria é calculado o seu cosseno do espaço euclidiano e confrontado com o cosseno do documento-mensagem para que se possa obter a medida de similaridade/proximidade entre a mensagem e cada categoria. Um vetor de categorias mais próximas é gerado no final deste processamento.

A última atividade do diagrama (atividade 9) é a pontuação dos especialistas, na qual são feitas inserções em duas tabelas do **FindYourHelp**, a primeira delas grava o *log* de mensagens postadas, onde um registro será criado nesta tabela identificando a mensagem e o número de palavras que pertencem a ela. Esta última informação será utilizada como forma de desempate pelo algoritmo de exibição dos especialistas, detalhado a seguir.

A segunda inserção é feita na própria tabela de *ranking* dos especialistas, onde o algoritmo inicialmente identifica se já foi criado algum registro para o participante e o tópico identificado e se, por ventura, ainda não existir registro na tabela contendo esta chave, o **FindYourHelp** criará um novo registro e atribuirá o valor 1 (um) à coluna de pontuação do mesmo, caso contrário, o registro encontrado será atualizado com o valor da pontuação já existente incrementado de um.

Entretanto, é necessário ressaltar que esta pontuação ao especialista só lhe é atribuída caso a mensagem postada alcance uma pontuação de similaridade em relação a uma categoria igual ou superior a um valor de referência fixado no código. A necessidade de estabelecer o ponto de corte foi percebida durante a execução do estudo de viabilidade da ferramenta e este será explicado com mais detalhes no capítulo que trata sobre o assunto.

Outro aspecto importante sobre esta atividade é que o **FindYourHelp** permite que os especialistas identificados automaticamente sejam validados pelo professor responsável pelo curso/grupo de pesquisa. Tal situação será detalhada no tópico que fala sobre a interação do usuário com o **FindYourHelp**.

5.4.3 Identificar especialistas para uma categoria escolhida

A identificação de especialistas acontece a partir da mesma tela que exibe a lista de categorias. Para esta funcionalidade, o objetivo é listar os participantes do fórum que possuem um grau de contribuição elevado no momento atual da consulta.

Esta tarefa requer uma atenção em relação ao número de mensagens categorizadas para o assunto selecionado ou os assuntos que o compõem, somando-se ainda a quantidade total de palavras analisadas de todas as mensagens já veiculadas por este especialista dentro do grupo. Este último item serve como um critério de desempate e ordenação entre os especialistas identificados. Este processo é revelado no diagrama abaixo.

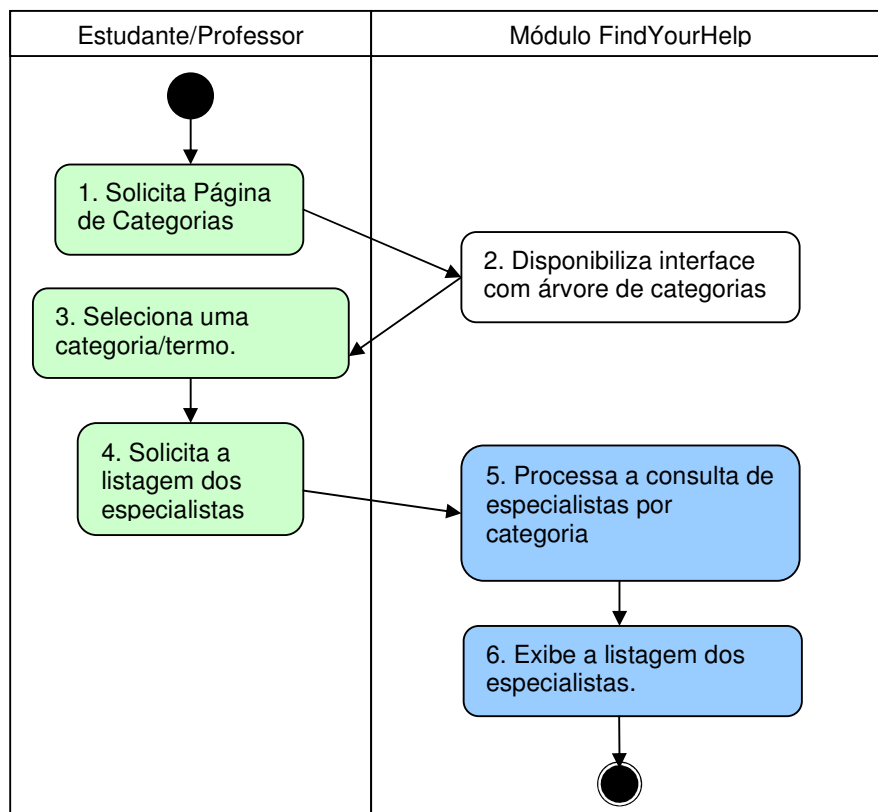


Figura 17 - Diagrama de atividades: Identificação e exibição de especialistas

O processo para listagem dos especialistas a partir de uma categoria ou termo existente começa semelhante ao de inserção de uma nova categoria, discutido anteriormente, ou seja, inicia com a solicitação da página de categorias, o **FindYourHelp** exibe então uma árvore de categorias (atividade 1) e traz a opção para que o professor/coordenador possa listar especialistas (atividade 2).

Deste ponto em diante, o usuário do sistema poderá optar pela listagem dos especialistas selecionando opcionalmente uma categoria para que esta seja definida como assunto central da consulta a ser gerada (atividade 3), considerando que nenhuma seja selecionada, a categoria raiz é utilizada como parâmetro.

A continuação da execução se dá com o processamento da consulta dos especialistas, partindo do assunto central escolhido na hierarquia de categorias (atividade 4). Neste momento é feita uma consulta a duas tabelas básicas do módulo **FindYourHelp**, a primeira será a tabela de *ranking* de especialistas e a segunda uma tabela de *log* que grava a quantidade de palavras digitadas para cada mensagem postada pelo participante (atividade 5). Em seguida, exibe-se uma listagem dos participantes dividida em 3 grupos de colaboração (A, B e C) (atividade 6) conforme já descritos anteriormente neste capítulo.

5.5 INTERAÇÃO COM O USUÁRIO DO AVA MOODLE

A interação do módulo **FindYourHelp** com o usuário do ambiente Moodle só acontece em três situações básicas, quais sejam: quando o usuário vai fazer a manutenção da árvore de categorias/tópicos, quando o usuário precisa buscar um especialista sobre determinado assunto e no momento em que se quer efetuar uma validação dos participantes identificados automaticamente como especialistas pela ferramenta.

É necessário ressaltar que, dentre essas atividades, a primeira e a terceira são disponibilizadas apenas para usuários do ambiente com o perfil de professor ou administrador.

5.5.1 Configurando a árvore de categorias

Uma das funcionalidades imprescindíveis para um bom funcionamento do algoritmo de categorização de mensagens é a tela que permite a configuração das categorias e tópicos a serem analisados.

A figura a 18 ilustra como esta tela é exibida para o usuário com perfil de professor ou administrador.

Categories



Figura 18 - Tela de exibição da árvore de categorias do módulo FindYourHelp

Conforme pode ser observado, esta tela possui a árvore de categorias listadas à esquerda e cinco botões de ações relativas a esta árvore. As categorias são representadas por um ícone amarelo, representando uma pasta, tal qual um diretório, e os tópicos que as compõem são representados por um ícone verde.

Dentre os botões exibidos, os quatro primeiros dizem respeito a ações que se ligam diretamente à tabela de hierarquia de categorias, já o último diz respeito à funcionalidade de exibição dos especialistas, e será detalhada mais adiante.

Os botões 1, 2 e 3 da figura 18, quando acionados, exibem a tela representada pela figura 19.

Category/Topic Name*

Category/Topic Description

There are required fields in this form marked*.

Figura 19 - Tela de inserção/edição de categorias/tópicos do módulo FindYourHelp.

Esta tela possui um funcionamento simples, contando com dois campos de texto e dois botões, onde o primeiro campo é o único de preenchimento obrigatório. Ao acionar o botão “*save changes*”, a tela efetuará a gravação dos dados na base de dados do módulo **FindYourHelp**, caso o botão “*cancel*” seja acionado, a tela anterior é exibida novamente. Além disso, a tela possui um comportamento distinto, quando seu botão “*save changes*” é acionado, de acordo com o botão clicado na tela de listagem de categorias.

Caso a tela tenha sido chamada a partir do primeiro botão da tela anterior (“*add new root category*”), uma nova categoria será inserida como raiz para subcategorias.

Caso o botão acionado tenha sido o segundo (“*add new child category or topic*”), uma categoria deve ser previamente selecionada na tela anterior e esta servirá como “categoria pai”, que será inserida pela tela de inserção/edição.

Caso o botão acionado seja o terceiro (“*edit category or topic*”), os dados da categoria ou assunto previamente selecionado aparecerão dentro dos campos da tela, dando a opção ao usuário de alterá-los e, posteriormente, salvar tal alteração na base de dados.

O quarto botão, quando acionado, exibirá uma tela de confirmação (figura 20) e, caso o usuário confirme, este apagará a categoria e suas subcategorias (caso existam), do contrário retornará à tela de listagem das categorias.

Delete Selected Category

Deleting Category/Topic. Confirm this operation?

Figura 20 - Tela de exclusão de categoria/tópico do módulo FindYourHelp

5.5.2 Listando os especialistas

O quinto botão da tela de listagem de categorias poderá ser acionado por qualquer participante ativo no ambiente Moodle. Ele produz uma listagem de participantes (*vide* figura 21) identificados pelo algoritmo do módulo como especialistas, a partir de algum assunto/categoria selecionado na tela anterior.

Experts (Desenvolvimento de Games)

A	✓ estudante 1 - estudante 2
C	✓ estudante 3
Legenda:	<p>A - Participantes que mais contribuíram com mensagens nos fóruns sobre este assunto até o presente momento!</p> <p>B - Participantes que contribuíram com mensagens nos fóruns sobre este assunto de forma moderada até o presente momento!</p> <p>C - Participantes que contribuíram em menor grau com mensagens nos fóruns sobre este assunto até o presente momento!</p>

Figura 21 - Lista de especialistas identificados pelo módulo FindYourHelp

A tela de listagem de especialistas apresenta os nomes de cada participante no formato de um *link* que aponta para a página do seu perfil (funcionalidade existente no próprio ambiente Moodle). Caso o botão “*Back to Category List*” seja clicado, a tela de listagem de categorias será exibida novamente.

Destaca-se que esta listagem exibe um ícone verde na forma de um “V” indicando que um determinado participante foi validado pelo professor como um especialista no assunto em questão, porém essa diferenciação visual acontece apenas para usuários do tipo professor conectados no sistema. Caso um usuário comum entre no sistema, não perceberá tal diferenciação entre as pessoas, portanto todos serão especialistas em sua visão.

Esta foi uma decisão de projeto, tendo em vista não incorrer em uma exposição desnecessária dos membros do grupo de discussão que não atendam aos pré-requisitos desta análise e, por conseqüência, evitar uma provável desmotivação dos estudantes envolvidos quanto à participação no processo de discussão.

Neste estudo, que tem como eixo a aprendizagem colaborativa (SILVA, 2003) é imprescindível assegurar que os sujeitos aprendam numa perspectiva horizontal, ajudando-se mutuamente. Em outras palavras, não é objetivo deste trabalho criar uma estrutura meritocrática, dentro de um espaço pedagógico, e que, desconsidere a subjetividade presente nos variados ritmos de aprendizagem, presentes em um espaço de interação como o AVA.

Faz-se necessário destacar que a ferramenta apresentada foi pensada com um caráter didático e uma intencionalidade que possa convergir para uma ação motivadora, na relação ensino-aprendizagem, do ponto de vista pedagógico. A validação do professor, em relação ao saber construído é um importante referencial para o discente, seja em espaços virtuais ou presenciais.

5.5.3 Validando os especialistas identificados

Visando obter uma credibilidade nos resultados apresentados pela ferramenta, criou-se uma possibilidade de validação humanamente assistida dos participantes identificados como especialistas.

Esta funcionalidade só poderá ser acessada por usuários com perfil de professor ou administrador e serve como um complemento às informações extraídas automaticamente pelo módulo **FindYourHelp**.

Uma das justificativas para esta funcionalidade se volta para a seguinte demanda docente, identificada na análise das entrevistas contidas no ANEXO A: professores de outros cursos, por exemplo, podem precisar contatar estudantes considerados competentes, do ponto de vista da elaboração de um conceito X, e que não pertençam ao seu quadro discente. Presume-se então, que eles irão recorrer à recomendação de colegas, que sugerirão quais são os estudantes que podem atender a esta demanda.

Considera-se ainda, no plano das suposições, que a validação docente é indispensável para a credibilidade das informações analisadas pela ferramenta, em

outras palavras, este professor demandante confiará mais no critério do professor que o aconselha, do que meramente na indicação automática gerada pelo algoritmo.

A figura 22 ilustra como o professor da disciplina poderá validar cada estudante apontado pelo módulo **FindYourHelp**.

Experts

	User Name	Status	Accepted	Denied
Projeto				
	estudante 1	A	<input checked="" type="radio"/>	<input type="radio"/>
	estudante 4	A	<input type="radio"/>	<input type="radio"/>
	estudante 2	C	<input checked="" type="radio"/>	<input type="radio"/>
Tipos Abstratos				
	estudante 3	A	<input type="radio"/>	<input checked="" type="radio"/>
	estudante 1	B	<input checked="" type="radio"/>	<input type="radio"/>

Figura 22 - Tela de validação de especialistas

A tela de validação conta com uma tabela onde será exibida uma listagem de estudantes que têm se destacado até o presente momento nos principais assuntos da hierarquia de categorias. Estes estudantes ficam agrupados pelos assuntos, para facilitar a visualização por parte dos professores.

Cada estudante é associado a uma letra, que representa o seu estado atual de contribuição nos fóruns sobre o assunto relacionado (coluna *Status*). Esta informação codificada pelas letras A, B e C, condiz com a da tela de listagem de especialistas, vista anteriormente.

O professor poderá então aceitar a indicação do algoritmo, validando assim sua percepção, para isto deverá selecionar o *radiobutton* da coluna *Accepted* correspondente ao estudante.

Caso contrário o professor poderá negar tal informação, selecionando o *radiobutton* da coluna *Denied* correspondente ao estudante. Tal escolha é mutuamente excludente, ou seja, só é possível marcar uma das duas colunas para cada participante listado.

Ao clicar no botão *confirm* os dados são persistidos em banco de dados e passam a vigorar na listagem de especialista, respeitando-se a situação do perfil de usuário conectado no momento.

5.6 CONSIDERAÇÕES FINAIS

Este capítulo apresentou a arquitetura e as funcionalidades da solução proposta pelo presente trabalho para a busca de especialistas em fóruns de discussão, intitulada **FindYourHelp**.

A solução foi projetada para ser um módulo acoplado a um ambiente virtual de aprendizagem, de forma que possa analisar postagens de mensagens diversas em fóruns de discussão através de um algoritmo de categorização de textos. O intuito é criar uma classificação de especialistas por categoria, sendo que estas devem ser previamente cadastradas pelo usuário responsável do grupo, geralmente o professor.

A idéia principal é que as mensagens sejam categorizadas e, no momento em que esta categorização é feita, ela pontua seu autor na tabela de *ranking* de especialistas. Em complemento, foi desenvolvida uma funcionalidade de validação humanamente assistida dos especialistas apontados pela ferramenta, visto que o algoritmo ainda está em desenvolvimento, portanto trata-se de um código ainda suscetível a falhas.

Utilizou-se a idéia do Modelo Espaço Vetorial para comparar os termos da mensagem com os termos que compõem cada categoria e as situações de sinonímia foram tratadas através do uso de um *thesaurus* em português (OPENTHESAURUS, 2009).

O próximo capítulo discorrerá sobre o estudo de caso desenvolvido para verificar a eficácia da solução desenvolvida.

6 FINDYOURHELP: ESTUDO DE VIABILIDADE DA FERRAMENTA

Diante do principal objetivo do presente trabalho, que se configura como a análise sobre o algoritmo de categorização de mensagens no apoio à busca por especialistas em um AVA, faz-se necessário a aplicação de um estudo de viabilidade da ferramenta aqui proposta (o módulo **FindYourHelp**) com o intuito de verificar e medir principalmente os seguintes aspectos do seu algoritmo:

- a) Quantidade de mensagens categorizadas corretamente de forma automática.
- b) Quantidade de mensagens descartadas corretamente pelo algoritmo.
- c) Pontuação de participantes confere com a percepção do professor responsável pelo grupo de estudo/pesquisa.

As sessões que se seguem demonstrarão o ambiente onde o estudo de viabilidade foi aplicado, em seguida descreverá como foi pensada a coleta dos dados necessários para a análise, bem como seus critérios de medição e validação e, por fim, uma análise dos resultados será demonstrada e detalhada.

6.1 O AMBIENTE DE APLICAÇÃO DO ESTUDO

Este estudo foi desenvolvido utilizando-se um ambiente de simulação contendo a versão 1.9.3 do AVA Moodle, disponível para download no sítio www.moodle.org. Para tanto o sistema foi instalado inicialmente em um computador de processador Intel Core 2 Duo, com 2.93Ghz de frequência e 2 gigabytes de memória RAM.

O foco de análise deste estudo se concentra em três disciplinas, as quais foram criadas no ambiente: Linguagem de Programação Orientada a Objetos I, Projeto Avançado de Sistemas II e Tecnologias Interativas Aplicadas à Educação.

É necessário ressaltar que todas essas disciplinas foram ministradas em caráter presencial, porém com uma estratégia de complementação dos estudos e da interação dos estudantes nos fóruns de pesquisas aqui analisados.

A identidade dos professores que participaram do presente estudo será preservada, sendo que aqui serão identificados como: Professor 1 (vinculado à disciplina Linguagem de Programação Orientada a Objetos I); Professor 2 (vinculado à disciplina Projeto Avançado de Sistemas II), e Professor 3 (vinculado à disciplina Tecnologias Interativas Aplicadas à educação).

É importante destacar que apenas o professor 3 não pôde responder à entrevista aplicada por este estudo, ficando a cargo de um dos estudantes da disciplina, que também é professor, a execução desta análise.

Uma entrevista semi-estruturada foi aplicada com cada um dos professores supracitados e maiores informações sobre suas formações e experiências, bem como um retorno contendo as impressões sobre a ferramenta aqui desenvolvida podem ser encontrados nos Apêndices deste trabalho.

6.2 A COLETA DOS DADOS

Os dados a serem coletados para análise se referem às mensagens postadas nos fóruns de discussão pertencentes às disciplinas supracitadas, conforme quadro abaixo:

Disciplina	Período
1 - Linguagem de Programação Orientada a Objetos I	18/08/2009 a 10/12/2009
2 - Projeto Avançado de Sistemas II	07/02/2009 a 18/06/2009
3 - Tecnologias Interativas Aplicadas à Educação	22/01/2004 a 31/01/2004

Quadro 5 - Disciplinas analisadas pelo estudo de viabilidade da ferramenta

Apenas a disciplina 1 já se encontrava em um sítio executando o AVA Moodle, portanto, para esta, foi feita uma importação dos dados para o ambiente de simulação deste trabalho através de uma cópia do banco de dados, as demais disciplinas possuíam seus fóruns de discussão em outros ambientes virtuais de colaboração, logo, foi necessário simular a postagem destas mensagens, uma a

uma na ordem em que ocorreram, dentro do ambiente de simulação instalado. Vale ressaltar que este último aspecto, trouxe uma percepção, durante a execução do estudo, sobre como se comporta o módulo **FindYourHelp** no momento da postagem de cada mensagem.

Para que o algoritmo processasse os dados da disciplina importada do ambiente Moodle, foi necessária a criação de uma página adicional ao ambiente virtual de simulação, construída na linguagem PHP, para que esta permitisse a escolha de um período de mensagens, onde se indique a data de início e data de término, a fim de que estas servissem como entrada de dados para o mesmo algoritmo de categorização e classificação de especialistas utilizado no momento de cada postagem. A figura 23 a seguir, exhibe um exemplo desta página.

The screenshot shows a web interface for 'Process Experts'. At the top, there is a green header with the title 'Process Experts' and a user login status 'You are logged in as [redacted] (Logout)'. Below the header is a breadcrumb trail 'Cursos de Informática > Process Experts'. The main content area is titled 'Process Experts' and contains two rows of date selection fields. The first row is for 'Initial Date' with values: 7, February, 2009, 00, 00, and a 'Disable' checkbox. The second row is for 'Final Date' with values: 18, June, 2009, 23, 55, and a 'Disable' checkbox. Below these fields is a 'Process Messages' button. At the bottom of the page, there is another user login status and a 'Home' link.

Figura 23 - Processamento de mensagens importadas do FindYourHelp

A página em questão é chamada a partir do próprio ambiente Moodle e, uma vez configurados os seus parâmetros, data inicial e data final, ela executará o algoritmo de categorização de mensagens e posterior identificação de especialistas de forma retroativa, respeitando o período definido, analisando assim suas respectivas mensagens.

Tais informações serão armazenadas aproveitando a estrutura já existente das tabelas do próprio módulo **FindYourHelp** para ranking de especialistas e log de mensagens categorizadas.

6.3 ANÁLISE: CRITÉRIOS E VALIDAÇÕES

A análise dos dados coletados foi desenvolvida mediante algumas questões de pesquisa levantadas previamente, tais como:

- a) Qual o percentual de acertos da categorização das mensagens por parte do algoritmo?
- b) Qual o percentual de acerto no descarte automático, por parte do algoritmo, das mensagens que não discutem os assuntos de interesse do professor da disciplina?
- c) Os especialistas apontados pelo algoritmo são reconhecidos pelo professor responsável pelo curso?

A partir destas questões, faz-se necessário analisar as mensagens postadas como forma de confrontar se o conteúdo veiculado na mensagem condiz com a(s) categoria(s) à(s) qual(is) foi(ram) associada(s). Para tanto, foi necessário obter a ajuda dos professores responsáveis por cada disciplina no momento desta análise.

Diante da quantidade de mensagens apresentadas pelos fóruns aqui analisados, foi tomada a decisão de se analisar todas as mensagens existentes, para trazer à presente pesquisa uma confiabilidade maior em seus resultados. Essa análise, portanto, responderá a primeira questão de pesquisa aqui levantada.

Para responder à segunda questão de pesquisa, o professor responsável pelo curso utilizará um recurso do módulo **FindYourHelp** para efetuar uma validação dos especialistas apontados pela ferramenta. Posteriormente uma análise será feita em cima das validações que este professor fizer a fim de confrontar quantos especialistas foram aceitos e quantos foram negados por ele frente ao total de especialistas apontados pela ferramenta.

Vale ressaltar que a entrevista semi-estruturada feita com os professores, citada anteriormente neste capítulo, entra como forma de complementar este

aspecto importante, que é a visão do professor sobre os resultados apontados pela ferramenta.

Portanto, visando um melhor aprofundamento sobre os resultados do estudo em questão, os critérios de análise adotados compreenderão os seguintes aspectos:

- a) Análise sobre as disciplinas escolhidas:
 - Nome da Disciplina e do Curso ao qual pertence
 - Instituição à qual está vinculada
 - Período de Atividade do Fórum
 - Quantidade de Participantes
 - Quantidade de Mensagens Postadas

- b) Análise sobre os docentes envolvidos:
 - Instituição
 - Formação
 - Disciplina que ministra
 - Tempo de experiência em Ensino
 - Tempo de experiência em EAD
 - *Feedback* sobre:
 - Experiência com o processo de montagem da hierarquia de termos
 - Identificação automática dos participantes pela ferramenta

- c) Análise sobre as postagens
 - Tipos de mensagem (somente texto, com elementos gráficos, links)
 - Quantidade de mensagens categorizadas corretamente
 - Quantidade de mensagens categorizadas com erro
 - Quantidade de mensagens descartadas corretamente
 - Quantidade de mensagens descartadas com erro
 - Participante(s) com maior número de postagens
 - Participante(s) com maior número de palavras escritas nas mensagens categorizadas

Sobre alguns conceitos acima citados, é válido esclarecer que a idéia de mensagens categorizadas corretamente aqui implica naquelas em que o texto evidencia que foram redigidas com o intuito claro de contribuir para a construção de conhecimento relacionado ao assunto identificado automaticamente pela ferramenta. Em contrapartida, mensagens categorizadas com erro implicam mensagens cujo texto não evidencia tal situação.

Complementando, as mensagens descartadas corretamente, são aquelas que em seu texto não tratam de assunto algum configurado na hierarquia de termos do ambiente Moodle da simulação, como por exemplo, mensagens para confirmar datas de trabalhos ou provas, bem como reportar ausências, reposições de aulas, ou perguntar quais assuntos foram abordados em aulas anteriores. Por fim, mensagens descartadas com erro, são aquelas em situação contrária a esta última descrita.

6.4 LIMITAÇÕES DO ESTUDO E AMEAÇAS ÀS VALIDAÇÕES

Outro ponto importante a ser pensado neste estudo foi referente à validade do experimento, a qual, segundo Travassos, Gurov e Amaral (2002), pode ser dividida em quatro tipos diferentes:

- a) Validade de Conclusão: aquela associada à habilidade de se chegar a uma conclusão sobre o resultado do experimento e a forma em que foi tratado. O tamanho do conjunto dos participantes, a confiabilidade das medidas e da implementação dos tratamentos devem ser observados com mais cuidado. Entretanto, a validade de conclusão do presente trabalho pode sofrer algumas ameaças também e essas serão melhor detalhadas no tópico que se segue.
- b) Validade Interna: é aquela que tem relação direta com os participantes do experimento, visto que estes podem influenciar os seus resultados. Problemas com esse tipo de validação podem decorrer de cansaços ou desânimos por parte dos participantes ou, em contrapartida, estes mesmos podem aprender ao longo do estudo. Considera-se que o presente estudo não apresenta esta ameaça, visto que as análises serão feitas sobre mensagens que já foram inseridas no fórum em um

período anterior e, portanto, não foram afetadas por problemas inerentes a esse tipo de validação.

- c) Validade de Construção: considera se o tratamento do experimento reflete bem a causa e se o resultado reflete bem o efeito. Comportamentos incorretos por parte dos participantes ou experimentador podem acarretar em problemas para o estudo. A validade de construção também pode ser entendida no presente estudo como um fator onde não se percebem problemas relacionados, pois se tratam de mensagens antigas, onde os comportamentos dos participantes/experimentador não influenciam nos resultados das análises.

- d) Validade Externa: aquela que define condições limitantes para que os resultados de um estudo possam ser generalizados para a prática industrial. Os problemas podem decorrer, por exemplo, da escolha de uma população de participantes não representativa para a indústria ou da aplicação do experimento em um dia/horário inadequado. A validade externa, por sua vez, consistiu um fator limitante para os resultados do presente estudo, como pode ser visto no tópico a seguir.

Os autores supracitados, consideram ainda que há uma prioridade de tipos de validade a ser definida de acordo com o experimento que se deseja executar. Para o presente estudo, por se tratar de um estudo aplicado (TRAVASSOS; GUROV; AMARAL, 2002), optou-se por priorizar os tipos de validade na seguinte ordem: interna, externa, construção e conclusão.

O presente estudo de viabilidade foi desenvolvido tendo em vista alguns fatores limitantes ao seu processo de execução e validação, dentre os quais é possível destacar:

1. A aplicação do estudo sobre mensagens trocadas por participantes de apenas três disciplinas: Tal situação, conforme já mencionado, se configura uma ameaça à validade de conclusão do presente trabalho.

2. A quantidade de mensagens analisadas foi um total de 325, somando-se as três disciplinas, e pode ser considerada aquém da quantidade desejada para uma análise mais aprofundada e conclusiva sobre o problema. Esta situação pode ser entendida como uma ameaça à validade externa do presente estudo, pois indica uma baixa representatividade frente às comunidades acadêmicas que se utilizam de recursos para ensino à distância existentes no mundo.

Visando um planejamento onde tais ameaças possam ser mitigadas ou eliminadas algumas respostas foram pensadas e serão aqui apresentadas tendo em vista serem colocadas em prática em trabalhos futuros relacionados.

Em relação à primeira ameaça, é necessário ressaltar que um trabalho posterior, utilizando um conjunto maior de participantes e uma análise de mensagens em tempo real de postagem nos fóruns, pode ser considerado como uma alternativa complementar ao presente estudo, visando sanar este aspecto e conduzindo a uma conclusão mais consistente sobre os resultados do trabalho.

Esta alternativa pode ainda mitigar a segunda ameaça anteriormente citada, pois pretende aumentar o conjunto de mensagens analisadas no estudo, porém não a elimina totalmente. Com a intenção de aperfeiçoar os efeitos desta resposta, foi pensada também como alternativa complementar, a análise de mensagens veiculadas por turmas provenientes de outras instituições de ensino e oriundas de áreas de formação distintas das que foram abordadas no presente estudo.

6.5 MONTAGEM DAS HIERARQUIAS DE CATEGORIAS

Para que o módulo **FindYourHelp** possa categorizar as mensagens do fórum corretamente, é necessário a alimentação da base de dados criando uma hierarquia de termos que definirá as principais categorias sobre as quais se deseja perceber uma relação com as mensagens postadas pelos estudantes participantes.

Cada uma das disciplinas citadas anteriormente como foco de análise para o presente trabalho deu origem a uma hierarquia de categorias e termos, as quais definem os principais assuntos ministrados pelos professores e estes, por sua vez, colaboraram com sua confecção e alimentação da base de dados.

A figura a seguir traz, como forma de exemplificação, a hierarquia de termos que compõe os principais assuntos da disciplina de Linguagem de Programação Orientada a Objetos I, definida em conjunto com o seu professor responsável.

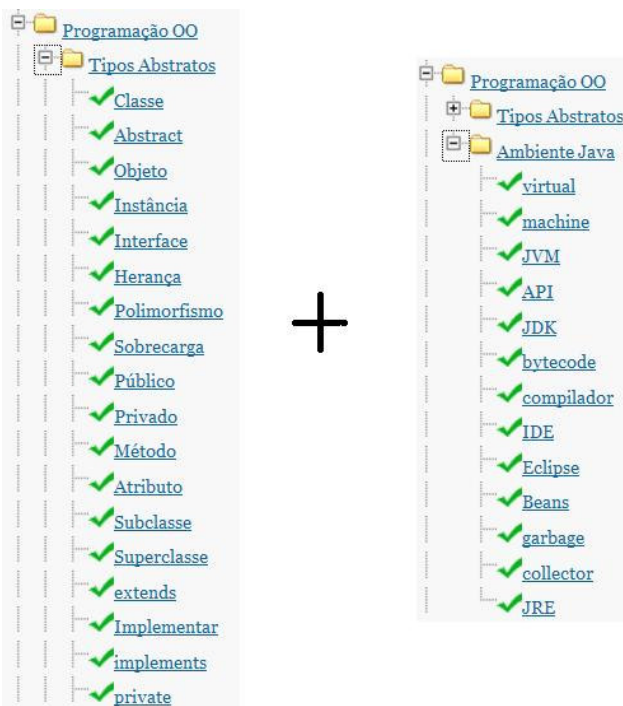


Figura 24 - Hierarquia de termos da disciplina de Linguagem de Programação Orientada a Objetos I

Como é possível perceber, a hierarquia acima traz em seu topo o assunto principal da disciplina, Programação Orientada a Objetos, e em seguida é subdividida em dois assuntos chaves da disciplina: Tipos Abstratos e Ambiente Java. Cada um desses assuntos é desmembrado então em termos os quais são citados com frequência em mensagens que falam sobre cada um deles. Por exemplo: uma mensagem com termos como “classe”, “objeto”, “instância”, “herança”, “método” e “atributo”, digitados em seu corpo, muito provavelmente é uma mensagem que fala sobre o assunto “Tipos Abstratos”.

6.6 ANÁLISE DOS RESULTADOS

Conforme já dito anteriormente, os resultados da presente pesquisa serão aqui divididos em três análises distintas: A análise sobre os docentes que

participaram do estudo; a análise sobre as disciplinas ministradas por estes docentes; e a análise sobre as mensagens postadas nestas disciplinas.

Para tanto, será adotado como ponto central a análise sobre a disciplina, partindo-se deste ponto para as demais informações investigadas.

É válido ressaltar que no decorrer deste estudo, as disciplinas foram submetidas ao algoritmo de categorização duas vezes, pois, durante a primeira execução, percebeu-se que o percentual de erro na categorização das mensagens estava alcançando até 20% em um dos casos, o que motivou uma análise e posterior alteração do algoritmo no trecho em que este decide quais categorias estão relacionadas com a mensagem postada.

Esta decisão foi baseada na observação de que o módulo estava relacionando mensagens a categorias com um nível de similaridade muito baixo. Percebeu-se, após a análise dos casos de erro, que as categorias com o valor de similaridade abaixo de 0,2 não condiziam com a mensagem analisada.

Desta forma, foi adicionada uma regra, a qual serviu como ponto e corte para mensagens com valor de similaridade muito baixo em relação a uma categoria, sendo esta regra montada a partir da seguinte condição: apenas mensagens com valor de similaridade maior ou igual a 0,2 serão categorizadas e pontuadas na tabela de *ranking* de especialistas.

O efeito desta modificação no algoritmo foi positivo, pois em todas as disciplinas analisadas houve um aumento no grau de acerto do algoritmo ao categorizar uma mensagem. É importante observar ainda que esta alteração não influenciou o descarte das mensagens. Este efeito será explicado durante a análise de cada uma das disciplinas a seguir.

6.6.1 Disciplina: Linguagem de Programação Orientada a Objetos I

Instituição: Centro Universitário da Bahia – FIB

Curso Vinculado: Sistemas de Informação

Quantidade de Participantes: 12

Quantidade Total de Mensagens: 32

Período de Atividade do Fórum: 10/08/2009 a 10/12/2009

Docente: Professor 1

Formação do Docente: Mestre em Sistemas e Computação em 2010 pela Faculdade Salvador – UNIFACS

Antes da alteração do algoritmo:

Mensagens Categorizadas Corretamente: 11

Mensagens Categorizadas com Erro: 2

Após a alteração do algoritmo:

Mensagens Categorizadas Corretamente: 12

Mensagens Categorizadas com Erro: 1

Mensagens Descartadas Corretamente: 19

Mensagens Descartadas com Erro: 0

Participante(s) com maior número de Postagens: Estudante 167 e Professor 1 com 10 postagens cada.

Participante(s) com maior número de palavras escritas nas mensagens categorizadas: Estudante 167 com 541 palavras

Especialistas Identificados:

- Assunto: Programação Orientada a Objetos:
 - Nível A: Estudante 167
 - Nível C: Estudante 342

Percebe-se nesta primeira disciplina analisada que a quantidade de mensagens foi pequena, com uma média de 0,27 mensagens por dia aproximadamente. A análise dos resultados desta disciplina revelou uma característica forte do algoritmo que é a capacidade de descartar mensagens, onde o acerto foi de 100% dos casos. Como exemplos, pode-se observar as mensagens abaixo:

Mensagem 1:

“Professor, envia por favor o material da última aula sobre classes abstratas e interface.”

Mensagem 2:

“Prof boa noite! Gostaria de confirmar o horário da av3 segunda dia 07/12”

Aqui é interessante perceber que apesar da mensagem 1 falar de termos como classes abstratas e *interface*, trata-se de uma solicitação ao professor e, portanto, não se configura como uma forma do participante contribuir com a discussão dentro do grupo.

O algoritmo de categorização do **FindYourHelp** identifica tal situação, na medida em que parte do pressuposto que o participante que inicia uma discussão no fórum geralmente é aquele que vai em busca da resposta e não aquele que possui a resposta, logo, estas mensagens são descartadas da contagem.

Já o aspecto da categorização das mensagens teve um percentual de erro de cerca de 8%, conforme figura a seguir.

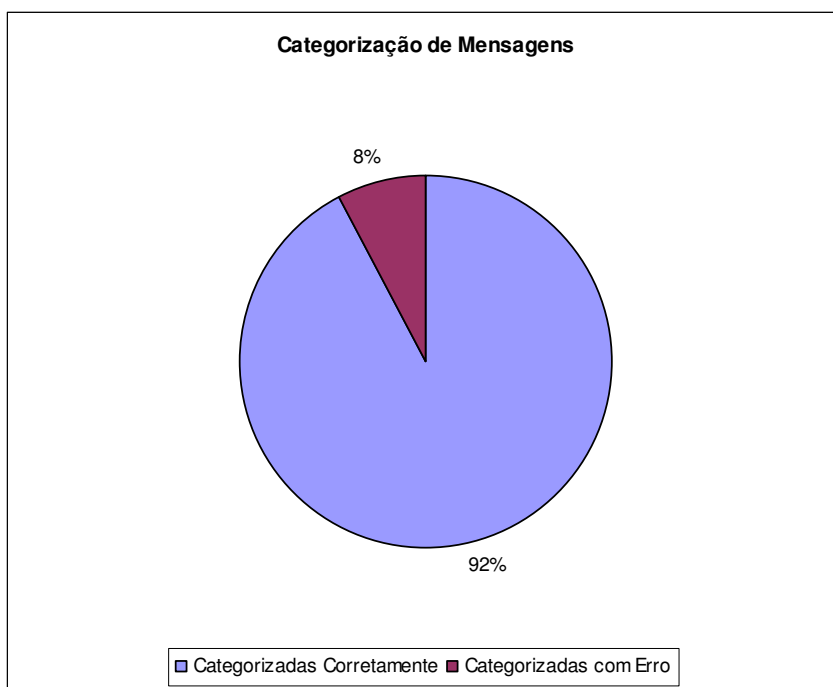


Figura 25 – Mensagens Categorizadas da Disciplina 1

Uma das mensagens categorizadas com erro pelo algoritmo antes da sua modificação e que foi corrigida após a aplicação da nova regra, pode ser conferida a seguir:

Mensagem 3:

“Aí galera, não está relacionado a LOOP, mas está relacionado a nossa área de tecnologia, fiquem a vontade para cadastrar e divulgar. http://www.programas2b.com.br/

Acho que vale a pena ter no □TTP□culo.“

A mensagem acima foi categorizada inicialmente com relação ao assunto “Programação OO Avançada”, pertencente a outra disciplina também cadastrada no ambiente de simulação. Em uma análise mais aprofundada, termos como “cadastrar”, “url” e “□TTP”, presentes no corpo da mensagem, podem ter influenciado a decisão do algoritmo, porém com a aplicação da regra de similaridade este problema desapareceu.

Observa-se ainda que o participante que mais postou mensagens no fórum é também considerado um especialista de nível A pela ferramenta no assunto central da disciplina.

6.6.2 Disciplina: Projeto Avançado de Sistemas II

Instituição: Centro Universitário da Bahia – FIB

Curso Vinculado: Sistemas de Informação

Quantidade de Participantes: 33

Quantidade Total de Mensagens: 76

Período de Atividade do Fórum: 07/02/2009 a 18/06/2009

Docente: Professor 2

Formação do Docente: Especialistas em Sistemas de Informação com Ênfase em Componentes Distribuídos e Web em 2000 pela Faculdade Ruy Barbosa – FRB

Antes da alteração do algoritmo:

Mensagens Categorizadas Corretamente: 32

Mensagens Categorizadas com Erro: 8

Após a alteração do algoritmo:

Mensagens Categorizadas Corretamente: 35

Mensagens Categorizadas com Erro: 5

Mensagens Descartadas Corretamente: 34

Mensagens Descartadas com Erro: 2

Participante(s) com maior número de Postagens: Estudante 56 com 18 postagens.

Participante(s) com maior número de palavras escritas nas mensagens categorizadas: Estudante 56 com 2653 palavras

Especialistas Identificados:

- Assunto: Projeto de Sistemas:
 - Nível A: Professor 2
 - Nível B: Estudante 56
- Assunto: Fundamentos do .NET
 - Nível A: Estudante 186 e Estudante 56
- Assunto: Acesso a Dados com .NET
 - Nível A: Professor 2
 - Nível B: Estudante 56
- Assunto: Sistemas Web com .NET
 - Nível A: Professor 2
 - Nível B: Estudante 56
- Assunto: Programação OO
 - Nível A: Professor 2
 - Nível B: Estudante 167 e Estudante 56
 - Nível C: Estudante 342

Percebe-se nesta segunda disciplina analisada que a quantidade de mensagens foi superior à quantidade da primeira disciplina (Linguagem de Programação Orientada a Objetos I), com uma média de 0,58 mensagens por dia aproximadamente. A análise dos resultados desta disciplina continuou a destacar uma característica forte do algoritmo que é a capacidade de descartar mensagens, onde o acerto desta feita foi em torno de 94% dos casos (figura 26).

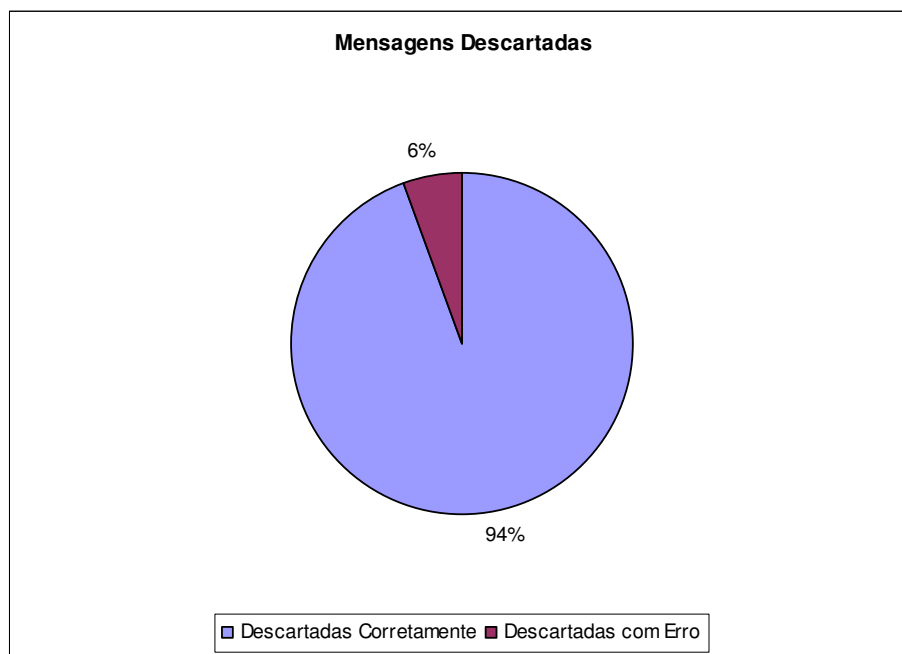


Figura 26 - Mensagens Descartadas da Disciplina 2

Neste caso, as duas mensagens descartadas foram as seguintes:

Mensagem 1:

“Será que já está utilizável? Achei os links e botei pra baixar vou testar junto com o Windows 7 que já tá instalado e rodando muito rápido. Segue link para download do Visual Studio 2010 e .NET Framework 4.0<<http://www.microsoft.com/downloads/details.aspx?FamilyId=922B4655-93D...>>”

Esta mensagem se trata de um *link* para baixar ferramentas importantes para a disciplina. Neste caso o algoritmo ignorou a mensagem erroneamente por se configurar a situação, já descrita anteriormente, em que o estudante estava abrindo uma discussão pela primeira vez.

Mensagem 2:

“fiz desse jeito, não sei se está correto, ainda vou confirmar com o professor, mas aqui insere no banco.

Tela

```
protected void Button1_Click(object sender, EventArgs e) {
String endereco = txtEndereco.Text;
etc
etc
CamadaFacade.Facade.InserirImovel(endereco,
complemntno,id_cidade,bairro,id_tipo,valor,obs, situacao);
}
```

```
public static void InserirImovel(String endereco, String complemento,
int id_cidade, String bairro, int id_tipo, float valor, String obs, int situacao){
new CamadaRN.ImovelRN().Inserir(endereco, complemento, id_cidade,
bairro, id_tipo, valor, obs, situacao); }
```

A situação acima descrita evidencia um problema do algoritmo de categorização com a interpretação de mensagens contendo código fonte, pois trechos como “CamadaFacade.Facade...” poderiam ser pistas para o algoritmo que sinalizassem que o autor da mensagem está discutindo o assunto “Padrões de Projeto”, porém percebe-se que o código não foi interpretado corretamente.

Tal situação pode ser explicada tendo em vista o momento em que o algoritmo retira os números, sinais de pontuação e caracteres especiais do texto da mensagem a ser processado. Neste exemplo, para um funcionamento correto do algoritmo, o sinal de ponto existente entre os termos “CamadaFacade” e “Facade” deveria ser suprimido e estes deveriam ser separados e espaços distintos, no vetor de termos da mensagem gerado.

Observa-se ainda que o percentual de categorização das mensagens com erro subiu para 13% nesta nova disciplina (figura 27).

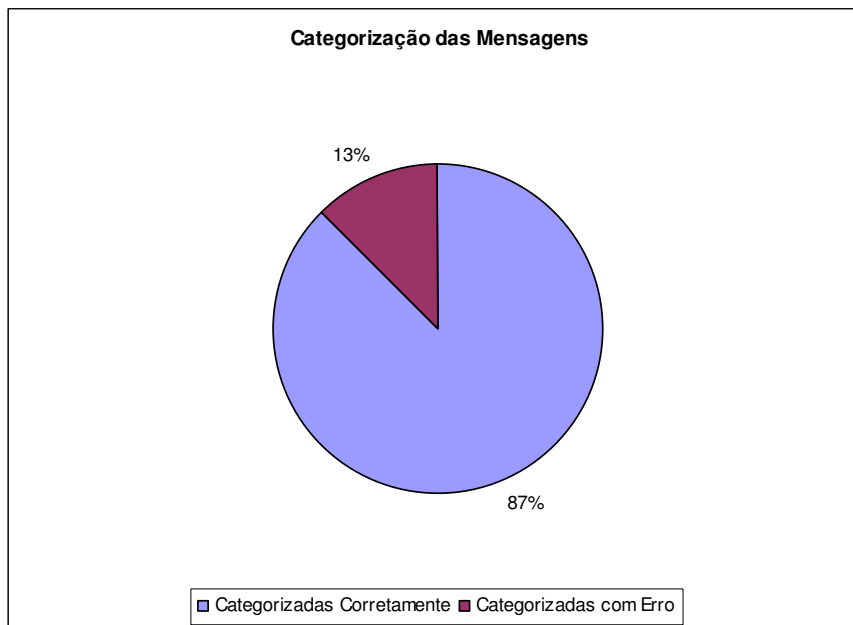


Figura 27 - Mensagens Categorizadas da Disciplina 2

É importante observar, entretanto, o seguinte: algumas mensagens desta análise, dentre as quais identificou-se uma categorização incorreta, na realidade foram categorizadas como pertencentes a alguns assuntos de forma correta e erroneamente em relação a outros, o que pode explicar este aumento de 8% para 13% nos erros de categorização do algoritmo. Observa-se tal situação no exemplo que se segue:

Mensagem:

“aproveitando o e-mail, nós estamos com várias dúvidas. As principais delas são:

- Conversão entre tipos de dados. (estamos encontrando dificuldade em relação ao parse) Por exemplo as conversões dos tipos int, char e datetime vem sendo um problema. Nós estamos testando todos os campos com String e mesmo assim vem ocorrendo erros após o submit. Acredito que esse problema também venha influenciando no meu delete do GridView; Como tratar essas exceções? Alguma dica?

- AutoPostBack, gostaríamos de carregar as informações de dentro de um

Banco de dados e exibir na tela para que houvesse um carregamento diferente no outro DropDownList. Como funciona essa sincronia? Para que o item selecionado no ddlOrigem não seja exibido no ddlDestino?”

Nesta situação o algoritmo categorizou a mensagem corretamente para o assunto “Tipos Abstratos”, porém relacionou a mensagem também com o assunto “Projeto de Games” de forma errada, neste caso notou-se que algumas palavras como: “dificuldade”, “problema”, “dica” e “sincronia” influenciaram o algoritmo a fazer uma relação errada, já que estes são termos vinculados ao assunto em questão.

Nesta disciplina, observa-se também uma maior participação do professor, o que fez com que o **FindYourHelp** o identificasse como participante especialista em muitos assuntos, porém alguns estudantes se destacaram também, dentre eles o que mais aparece relacionado nos assuntos é o estudante 56, o qual foi reconhecido pelo professor da disciplina durante a sua análise sobre os resultados (ver entrevista no ANEXO B).

Um dado complementar a ser percebido é que o referido estudante foi o que mais postou mensagens dentro do grupo e o que mais escreveu palavras em suas mensagens também, porém isso não garantiu que ele fosse classificado como Nível A em todos os assuntos, o que evidencia que o algoritmo utilizado não se baseia apenas na quantidade das postagens dos usuários.

6.6.3 Disciplina: Tecnologias Interativas Aplicadas à Educação

Instituição: Universidade Católica do Salvador - UCSAL

Curso Vinculado: Metodologia e Didática do Ensino Superior

Quantidade de Participantes: 31

Quantidade Total de Mensagens: 217

Período de Atividade do Fórum: 22/01/2004 a 30/01/2004

Docente: Professor 3

Formação do Docente: Doutora em Educação pela Universidade Federal da Bahia - UFBA em 2004

Antes da alteração do algoritmo:

Mensagens Categorizadas Corretamente: 118

Mensagens Categorizadas com Erro: 13

Após a alteração do algoritmo:

Mensagens Categorizadas Corretamente: 125

Mensagens Categorizadas com Erro: 6

Mensagens Descartadas Corretamente: 80

Mensagens Descartadas com Erro: 3

Participante(s) com maior número de Postagens: Estudante 360 com 27 postagens

Participante(s) com maior número de palavras escritas nas mensagens categorizadas: Estudante 360 com 3117 palavras

Especialistas Identificados:

- Assunto: Tecnologias da Informação e da Comunicação
 - Nível A: Estudante 360
- Assunto: Políticas de informatização, Programas Governamentais, Inclusão e Exclusão Digital
 - Nível A: Estudante 363, Estudante 357, Estudante 360 e Estudante 356
- Assunto: Interatividade, Interação e Hipertextualidade
 - Nível A: Professor 3
- Assunto: Pedagogia da Diferença
 - Nível A: Estudante 359
- Assunto: Escola e Produção de Conhecimento
 - Nível A: Estudante 381, Estudante 359, Estudante 372, Estudante 357 e Estudante 358
 - Nível B: Estudante 360

Percebe-se nesta terceira disciplina analisada que a quantidade de mensagens foi superior ao somatório das outras duas disciplinas (214 mensagens

postadas), bem como a média de mensagens por dia aumentou significativamente para 23,78.

Este último fato foi explicado pelo professor responsável pela análise da disciplina, o qual afirmou que um dos objetivos da mesma era que o grupo envolvido pudesse vivenciar na prática o uso das tecnologias de informação e comunicação no dia a dia de uma turma de pós-graduação.

A análise dos resultados desta disciplina confirmou em princípio a característica a capacidade do algoritmo de descartar mensagens, onde o acerto desta feita foi em torno de 96% dos casos (figura 28).

Durante esta análise, percebeu-se que as mensagens descartadas com erro pelo algoritmo foram decorrentes de uma baixa pontuação encontrada pelo método TFIDF devido a uma característica peculiar do grupo de participantes desta disciplina: eles escreveram um maior número de palavras por mensagens frente aos estudantes das outras disciplinas.

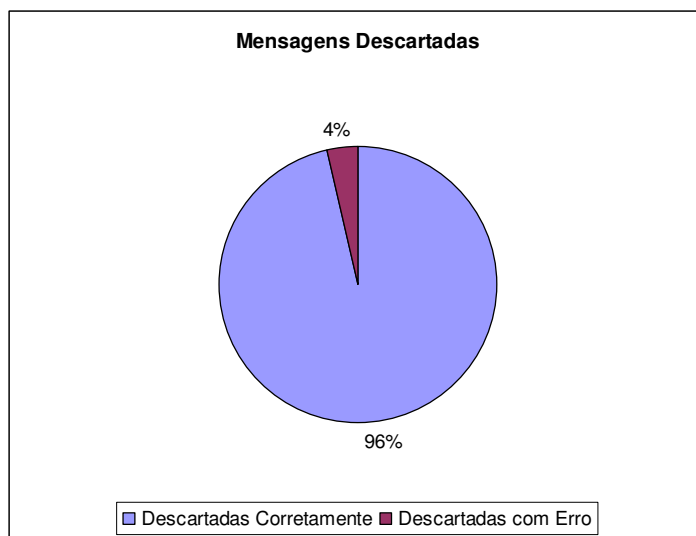


Figura 28 - Mensagens descartadas da Disciplina 3

Em uma observação mais aprofundada, após alguns testes, percebeu-se que a abordagem TFIDF do algoritmo sofre uma alteração para menor na pontuação atribuída a uma mensagem, que contenha a mesma quantidade de palavras relacionadas a uma categoria em comparação com outra mensagem, porém com maior número de palavras não relacionadas à categoria. Em outras palavras, percebe-se que, para o algoritmo, um participante tido como prolixo tem sua mensagem

pontuada com um valor menor do que um participante que se expresse com maior objetividade sobre o mesmo assunto.

Sobre a categorização das mensagens, esta disciplina obteve o resultado mais expressivo, sendo 95% das mensagens categorizadas corretamente. Os problemas encontrados nesta situação foram decorrentes dos mesmos fatores já citados nas outras disciplinas.

Novamente pode ser observada a situação em que o estudante que mais postou mensagens não foi classificado como nível A em todos os assuntos. Um fato não previsto sobre esta disciplina foi que os estudantes postavam muitas mensagens que representavam as discussões e idéias de um grupo inteiro, portanto, na interpretação do algoritmo, a pontuação foi atribuída apenas para o usuário que efetuou a postagem, o que, de certa forma, produziu algumas distorções nos resultados apresentados sobre quem realmente seria considerado especialista dentro do grupo.

O quadro comparativo abaixo resume as principais informações sobre as disciplinas aqui analisadas:

Disciplina	1) Linguagem de Programação Orientada a Objetos I	2) Projeto Avançado de Sistemas II	3) Tecnologias Interativas Aplicadas à Educação
Tipo	Presencial	Presencial	Presencial
Tema Central	Linguagem de Programação	Projeto de Sistemas	Educação e Tecnologias
Total de Participantes	12	33	31
Total de Mensagens	32	76	217
Ano	2009	2009	2004
Duração	Cerca de 4 meses	Cerca de 4 meses	9 dias
Mensagens Categorizadas	OK = 12 (92%) ERRO = 1 (8%)	OK = 35 (87%) ERRO = 5 (13%)	OK = 125 (95%) ERRO = 6 (5%)
Mensagens Descartadas	OK = 19 (100%)	OK = 34 (94%) ERRO = 2 (6%)	OK = 80 (96%) ERRO = 3 (4%)
Docente concorda com especialistas identificados?	SIM, TOTALMENTE	SIM, TOTALMENTE	SIM, TOTALMENTE

Quadro 6 - Comparativo entre os resultados das disciplinas analisadas

Este quadro revela ainda algumas informações importantes para esta análise, tais como a diferença de cinco anos entre a disciplina 3 e as demais, o que leva uma hipótese sobre o comportamento dos estudantes envolvidos nesta disciplina, de que estes estavam provavelmente motivados por uma novidade tecnológica no período em que ela aconteceu. Tal situação fica mais evidenciada se for percebida a diferença de tempo de duração entre elas, onde mais uma vez a disciplina 3 revelou uma intensidade muito grande de mensagens veiculadas em um curto espaço de tempo, frente as demais.

Um ponto positivo para os resultados do algoritmo encontra-se na concordância de todos os docentes sobre as indicações de especialistas por parte da ferramenta. Dentre eles, apenas o docente da disciplina 3 ficou em dúvida sobre um dos participantes apontados antes que o algoritmo de categorização fosse alterado com a regra do valor de similaridade mínimo, mas concordou totalmente após a modificação no mesmo, o que evidencia uma eficiência da solução implementada neste aspecto da análise.

6.7 CONSIDERAÇÕES FINAIS

O presente capítulo apresentou o estudo de viabilidade da ferramenta **FindYourHelp**. Para tanto, foram utilizadas as mensagens provenientes de três grupos de discussão referentes a três disciplinas distintas, todas ofertadas em caráter presencial em suas instituições. Tais disciplinas utilizaram os fóruns virtuais de forma a complementar seus recursos pedagógicos.

Para executar o estudo, as mensagens foram importadas para o ambiente de testes descrito neste capítulo, com o intuito de simular a inserção de cada mensagem postada por cada participante das disciplinas. Foram convidados também três professores para que atuassem como usuários da ferramenta a fim de criar uma hierarquia de categorias para suas respectivas disciplinas, bem como para validar os resultados apontados pelo **FindYourHelp** ao final do processo.

Os resultados obtidos e opiniões dos docentes participantes deste estudo trazem a percepção de que o módulo criado para o AVA Moodle atende satisfatoriamente ao objetivo para o qual foi proposto, o qual consiste em apoiar

professores e coordenadores de curso na identificação de participantes especialistas em assuntos específicos tratados dentro de um fórum de discussão no Moodle.

No próximo capítulo serão apresentadas as contribuições desta pesquisa, as perspectivas de trabalhos futuros relacionados, bem como as limitações vivenciadas.

7 CONCLUSÃO

Constantes são as transformações pelas quais passa a Sociedade Contemporânea e dentre elas, pode-se destacar o surgimento da Sociedade do Conhecimento. É sabido que o conhecimento recentemente posicionou-se no centro das ambições humanas. Seja no âmbito empresarial ou acadêmico, os “detentores” de conhecimento são pessoas valorizadas em suas instituições.

Estudos sobre Gestão do Conhecimento apontam que para essas pessoas contribuírem ao máximo com o ambiente em que estão inseridas, deve-se buscar formas de socialização deste conhecimento com as demais, evitando que este fique estagnado com o seu detentor.

Com o surgimento de ambientes virtuais de aprendizagem (AVA) apoiando a colaboração em grupos de discussão em diversas Instituições de Ensino Superior (IES), percebeu-se a criação de uma grande quantidade de informações que circulam e são armazenadas nestas grandes bases de dados acadêmicas.

Durante essa pesquisa, foi percebido que os professores que utilizam tais ferramentas, buscam formas de mobilizar a participação de seus estudantes e alguns gostariam de avaliar a participação destes integrantes de forma mais detalhada, porém a maioria dos AVA aqui pesquisada não dispõe de recursos que possam afirmar, por exemplo: qual estudante participa mais quando a discussão se remete ao assunto X?

Em grupos de discussões suportados por ferramentas como fóruns virtuais, geralmente se percebe a existência de participantes com maiores habilidades e competências em determinados assuntos. Tais pessoas podem ser consideradas como especialistas no assunto em questão perante as outras que participam do mesmo grupo.

Encontrar rapidamente estas pessoas se configura uma demanda de grande importância não só para os professores de uma disciplina, como também para os demais integrantes do grupo que estejam diante de situações ou problemas relacionados ao assunto que um determinado especialista domina.

Neste sentido, o presente trabalho apresentou uma proposta para tratar esta situação em um AVA muito utilizado por IES, chamado Moodle. Para tanto um módulo, intitulado FindYourHelp, foi desenvolvido e adicionado ao ambiente em

questão de forma a interceptar as mensagens postadas pelos usuários da ferramenta de fórum de discussão.

O módulo utiliza técnicas de mineração de textos tais como a montagem de um vetor de termos significativos para cada mensagem postada, bem como a categorização automática de mensagens diante da similaridade destas em relação a uma hierarquia de termos previamente cadastrada pelos professores.

A ferramenta foi analisada em conjunto com três docentes de disciplinas distintas, os quais montaram uma hierarquia de categorias/assuntos que representasse cada uma de suas disciplinas. Eles avaliaram também os resultados apontados pelo **FindYourHelp** no que tange os participantes identificados automaticamente como especialistas para cada assunto central da disciplina.

Dentre as disciplinas analisadas, percebeu-se que uma delas claramente se propôs ao uso do fórum como um recurso principal no cotidiano da turma envolvida e esta contribuiu para oferecer um campo de análise consistente para o estudo de viabilidade da ferramenta. As outras duas disciplinas utilizaram o recurso do fórum como um apoio ao ensino presencial, e nelas percebeu-se que a utilização deste para o fim de discussão dos assuntos foi menos intensa por parte de seus integrantes.

Nesse contexto, percebeu-se que a solução proposta neste trabalho atende satisfatoriamente ao objetivo para o qual foi desenvolvida, bem como foram levantadas sugestões para possíveis melhorias futuras.

7.1 CONTRIBUIÇÕES

A partir da análise descrita neste capítulo, a principal contribuição percebida neste trabalho consiste em: disponibilizar para a comunidade acadêmica uma alternativa à busca de participantes especialistas em assuntos determinados dentro de um AVA, de código fonte aberto, através da análise das postagens efetuadas em seus fóruns de discussão.

Dentre as principais contribuições deste trabalho é possível destacar:

- a) Estudo comparativo de ambientes virtuais e aprendizagem

- b) Proposta de uma solução, intitulada **FindYourHelp**, para mineração das mensagens postadas em fóruns de discussão, em forma de um *plugin* adicional ao ambiente Moodle, capaz de identificar automaticamente os participantes mais ativos que discutem assuntos específicos das disciplinas cadastradas
- c) Desenvolvimento de métodos de categorização de mensagens de acordo com o grau de similaridade existente entre estas e as categorias definidas pelos docentes de cada disciplina, com ênfase na técnica de distância de cosseno.
- d) Análise da aplicabilidade de um thesaurus disponível para português (OPENTHESAUROS, 2009) de forma a reduzir a dimensão do conjunto de termos mais significativos das mensagens dos fóruns
- e) Análise do uso de uma funcionalidade humanamente assistida como forma de incrementar a credibilidade dos resultados apontados pela solução desenvolvida
- f) Desenvolvimento de um estudo de viabilidade para validação da ferramenta implementada.

A metodologia adotada no desenvolvimento desta dissertação possibilitou também a aplicação de pesquisas bibliográficas e a revisão de literatura sobre os temas de gestão do conhecimento, buscas de especialistas, ambientes virtuais de aprendizagem, mineração de textos e categorização automática de textos. Tal estratégia foi de suma importância para definir as características da solução construída.

7.2 TRABALHOS FUTUROS

O estudo de viabilidade da ferramenta **FindYourHelp** permitiu vislumbrar alguns pontos com necessidade de melhorias, possíveis de serem implementadas em trabalhos futuros.

Dentre as possíveis melhorias, a principal a ser destacada é investigar a utilização de outras técnicas de categorização de textos para confrontar a eficiência do método aqui proposto. Além desta sugestão, os seguintes trabalhos futuros podem ser desenvolvidos como forma de evoluir a solução proposta:

- a) Melhorias referentes à interface da ferramenta com o usuário, para permitir uma melhor interação no momento de definir as categorias/assuntos relacionados à disciplina. Para tanto, recomenda-se a utilização de recursos avançados de desenvolvimento de interface visual de aplicações web, tais como o AJAX.
- b) Análise através de um experimento controlado e executado em tempo real, compreendendo uma ou mais disciplinas que estejam em andamento, para avaliar questões como o comportamento dos estudantes quando cientes de que uma forma de monitoramento das mensagens estará ativa ao executarem suas postagens nos fóruns, bem como questões relacionadas a motivação ou desmotivação quando percebem a ferramenta identificando-os ou a seus colegas como especialistas.
- c) Criação de mecanismos para manipulação do Thesaurus existente de forma a permitir maior suporte a línguas estrangeiras e proporcionar melhor desempenho ao algoritmo de categorização. Tal sugestão baseia-se em situações detectadas durante o estudo de viabilidade da ferramenta, as quais revelaram que disciplinas como Linguagens de Programação Orientada a Objetos I se utilizam constantemente de palavras em inglês que, nestes casos, são consideradas chaves para a detecção do algoritmo de categorização. No entanto as mensagens poderiam escrevê-las utilizando termos de mesmo significado escritos em língua portuguesa, o que forçou o cadastramento desnecessário de dois termos referentes ao mesmo significado na hierarquia de categorias.

- d) Estudo de viabilidade da aplicação de técnicas de reconhecimento semântico nos textos das mensagens postadas nos fóruns, tal como a indexação semântica latente (LSI), visando obter uma maior eficácia na categorização promovida pelo algoritmo.

7.3 LIMITAÇÕES E CONSIDERAÇÕES FINAIS

Esta pesquisa propôs uma solução de busca por especialistas em um AVA, de código fonte aberto, desenvolvido em PHP. Esta situação impôs uma restrição técnica importante ao desempenho do algoritmo desenvolvido, pois a linguagem PHP não possui suporte nativo ao processamento paralelo de linhas de execução (*multithreading*), o que possivelmente produziria um ganho de performance considerável em sua execução caso pudesse ser implementado no código.

Outro fator limitante ao trabalho foi a impossibilidade de analisar uma quantidade maior de IES e maior diversidade de disciplinas no Moodle, pois nem todas as instituições o utilizam nem possibilitam o acesso a estas informações. Neste caso o tempo para a execução da pesquisa seria muito estendido na tentativa de encontrar outras instituições para incrementar o presente estudo.

Por fim, a solução proposta neste trabalho obteve uma aceitação considerada satisfatória por parte dos docentes nele envolvidos, o que indica que a solução pode ser considerada aplicável em novas situações de análise e em outras instituições.

REFERÊNCIAS

ADOBE. **Adobe flash platform**. Disponível em: <<http://www.adobe.com/flashplatform/>>. Acesso em: 16 dez. 2009.

BALOG, Krisztian; RIJKE, Maarten. Determining expert profiles (with an application to expert finding). In: INTERNATIONAL JOINT CONFERENCE ON ARTIFICIAL INTELLIGENCE, 20., 2007, Hyderabad/India. **Anais...** Hyderabad/India, 2007. Disponível em: <<http://www.ijcai.org/papers07/Papers/IJCAI07-427.pdf>>. Acesso em: 7 jan. 2009.

BASTOS, Valéria M. **Ambiente de descoberta de conhecimento na web para a língua portuguesa**. Rio de Janeiro: UFRJ, 2006.

BECERRA-FERNANDEZ, Irma. Searching for experts on the web: a review of contemporary expertise locator systems. **ACM Transactions on Internet Technology (TOIT)**, v. 6, n. 4, p.333-335, nov. 2006.

BLACKBOARD. **Blackboard**. Disponível em: <<http://www.blackboard.com/>>. Acesso em: 25 nov. 2008.

CARVALHO, Rodrigo, B. de. **Aplicações de software de gestão do conhecimento**: tipologia e usos. Belo Horizonte: Escola de Ciência da Informação - UFMG, 2000.

COELHO, Alexandre R.; ORENGO, Viviane M.; BURIOL, Luciana S. **Removedor de sufixos da língua portuguesa**: RSLP. 2007. Disponível em: <<http://www.inf.ufrgs.br/~arcoelho/rspl/html/index.html>>. Acesso em: 20 nov. 2009.

DAVENPORT, Thomas H.; PRUSAK, Laurence. **Conhecimento empresarial**: como as organizações gerenciam seu capital intelectual. 13. ed. Rio de Janeiro: Elsevier, 2003.

DIAS, Cláudia A. Portal corporativo: conceitos e características. **Ci. Inf.**, v. 30, n.1, jan./abr. 2001. Disponível em: <http://www.scielo.br/scielo.php?pid=S0100-19652001000100007&script=sci_arttext&tlng=es>. Acesso em: 25 jun. 2008.

DOKEOS. **Dokeos e-learning**. Disponível em: <<http://www.dokeos.com/>>. Acesso em: 25 nov. 2008.

EDUWEB. **EduWeb**. Disponível em: <<http://www.eduweb.com.br/home.asp>>. Acesso em: 25 set. 2008.

ELLIS, C. A.; GIBBS, S. J.; REIN, G. L. Groupware: some issues and experiences. **Communications Of The ACM**, v.134, n.1, jan. 1991.

FARIA, Elaine T. (Org.) et al. **Educação presencial e virtual: espaços complementares essenciais na escola e na empresa.** Porto Alegre, RS: EDIPURS, 2006.

FAYYAD, Usama, PIATETSKY-SHAPIRO, Gregory, SMITH, Padhraic. The kdd process for extracting useful knowledge from volumes of data. **Communications of the ACM**, v. 39, n.11, p.27-34, nov. 1996a.

_____. From data mining to knowledge discovery in databases. **AI Magazine**, 1996b

FIALHO, Francisco A. P. **Gestão do conhecimento e aprendizagem: as estratégias competitivas da sociedade pós-industrial.** Florianópolis, SC: Visual Books, 2006

FILIPPO, Denise; FUKS, Hugo; LUCENA, Carlos J. P. AulaNetM: extensão do serviço de conferências do aulanet destinada a usuários de pdas. In: SIMPÓSIO BRASILEIRO DE INFORMÁTICA NA EDUCAÇÃO – SBIE, 16., 2005. Juiz de Fora, MG **Anais...** Juiz de Fora, MG, 2005.

GALLUPE, R. B. **Knowledge management systems: surveying the landscape.** Queen's University at Kingdom, 2000

GAMMA, Eric et al. **Padrões de projeto: soluções reutilizáveis de software orientado a objetos.** Tradução: Luiz A. Meirelles Salgado. Porto Alegre, RS: Bookman, 2000.

GEROSA, M.A. et al. Uma arquitetura para o desenvolvimento de ferramentas colaborativas para o ambiente de aprendizagem aulanet. In: SIMPÓSIO BRASILEIRO DE INFORMÁTICA NA EDUCAÇÃO – SBIE, 2004. Manaus. **Anais...** Manaus, 2004.

GEROSA, M.A. et al. O uso de uma arquitetura baseada em componentes para incrementar um serviço do ambiente aulanet. In: WORKSHOP DE DESENVOLVIMENTO BASEADO EM COMPONENTES – WDBC, 4., 2004, João Pessoa, PB. **Anais...** João Pessoa, PB, 2004.

JENNEX, Murray. Knowledge management: concepts, methodologies, tools and applications. **Information Science Reference**, v. 1, USA, 2008.

JUNG, Hanming et al. Finding topic-centric identified experts based on full text analysis. In: INTERNATIONAL EXPERTFINDER WORKSHOP, FEWS, 2., 2007, Busan – Korea. **Anais...** Busan – Korea, 2007.

KONCHADY, Manu. **Text mining application programming.** Boston: Charles River Media, 2006

LÉVY, Pierre. **Cibercultura.** São Paulo: Editora 34, 1999.

LINGUATECA. **Lista de palavras frequentes em português.** Disponível em: <<http://linguateca.di.uminho.pt/Paulo/stopwords/>>. Acesso em: 17 maio 2009.

MAGALHÃES, Cristiane, C. **Minerjur**: uma ferramenta para mineração de bases de jurisprudência. Salvador, BA: UNIFACS, 2008.

MATTOX, D.; MAYBURY, M.; MOREY, D. Enterprise expert and knowledge discovery. In: INTERNATIONAL CONFERENCE ON HUMAN-COMPUTER INTERACTION. 1999. **Anais...** 1999.

MAYBURY, Mark T. **Expert finding systems**. Mitre Technical Report, 2006

MOODLE. **Moodle sites**. Disponível em: <<http://moodle.org/sites/index.php?country=BR>>. Acesso em: 20 nov. 2008a.

_____. **História do moodle**. Disponível em: <http://docs.moodle.org/pt_br/História_do_Moodle>. Acesso em: 20 nov. 2008b.

_____. **Modules and plugins**. Disponível em: <<http://moodle.org/mod/data/view.php?id=6009>> Acesso em: 21 nov. 2008c.

MORIN, Edgard. **Os sete saberes necessários à educação do futuro**. 3. ed. São Paulo:Cortez; Brasília, DF: UNESCO, 2001.

MYSQL. **Mysql**: developer zone. Disponível em: <<http://dev.mysql.com>>. Acesso em: 16 dez. 2009.

NABER, Daniel. **Openthesaurus**: building a thesaurus with a web community. 2004 Disponível em: <<http://openthesaurus.caixamagica.pt/download/openthesaurus.pdf>>. Acesso em: 10 set. 2009.

NONAKA, Ikujiro; TAKEUCHI, Hirotaka. **Criação de conhecimento na empresa**: como as empresas japonesas geram a dinâmica da inovação. 18. ed. Rio de Janeiro, RJ: Elsevier, 1997.

OPENTHESAURUS. **Português thesaurus in text format**. Disponível em: <<http://openthesauruspt.caixamagica.pt>>. Acesso em: 29 set. 2009

ORACLE. **Sun developer network**: Applets. Disponível em: <<http://java.sun.com/applets/>>. Acesso em: 16 dez. 2009

PIATETSKY-SHAPIRO, Gregory et al. Is there a grand challenge or x-prize for data mining? In: ACM SIGKDD INTERNATIONAL CONFERENCE ON KNOWLEDGE DISCOVERY AND DATA MINING, 12., 2006. Philadelphia, Pennsylvania, USA, . **Proceedings...** Philadelphia, Pennsylvania, USA, 2006.

RESENDE, S. O. (Coord). **Sistemas inteligentes**: fundamentos e aplicações. Barueri, SP: Manole, 2005.

RICHARDS, Debbie; TAYLOR, Meredith; BUSCH, Peter. Expertise recommendation: a two-way knowledge communication channel. In: INTERNATIONAL CONFERENCE ON AUTONOMIC AND AUTONOMOUS SYSTEMS, 4., 2008. **Anais...** 2008.

ROBERTSON, S. Understanding inverse document frequency: on theoretical arguments for IDF. **Journal of Documentation** **60**, n. 5, p. 503-520, 2004.

SALTON G.; WONG, A.; YANG C. S. A vector space model for automatic indexing. **Communications of the ACM**, New York, v.18, 1975.

SANTOS, Marcos L.; SALVADOR, Laís do N. FindYourHelp: um módulo de busca por especialistas no ambiente Moodle. In: SIMPÓSIO BRASILEIRO DE INFORMÁTICA NA EDUCAÇÃO, 20., 2009. Florianópolis, SC, **Anais...** Florianópolis, SC, 2009.

SETZER, Valdemar W. Dado, informação, conhecimento e competência, **DataGramZero - Revista de Ciência da Informação**, n. 0, dez.1999. Disponível em: <http://www.dgz.org.br/dez99/F_I_art.htm>. Acesso em: 20 jul. 2008.

SILVA, Marco. Educação na cibercultura: o desafio comunicacional do professor presencial e online. **Revista da FAEEBA – Educação e Contemporaneidade**, Salvador, v. 12, n. 20, p. 261-271, jul./dez., 2003.

SOUCY, Pascal; MINEAU, Guy W. Beyond tfidf weighting for text categorization in the vector space model. In: INTERNATIONAL JOINT CONFERENCE ON ARTIFICIAL INTELLIGENCE (IJCAI), 19., 2005. **Proceedings...** 2005.

SUN, **Java ee at a glance**. Disponível em: <<http://java.sun.com/javaee/>> Acesso em: 19 set. 2008.

TELEDUC. **TelEduc**: Ensino à Distância. Disponível em: <<http://www.teleduc.org.br/>>. Acesso em: 27 nov. 2008.

TOFFLER, Alvin. **A terceira onda**. Rio de Janeiro: Record, 1993.

TRAVASSOS, Guilherme H.; GUROV, Dmytro; AMARAL, Edgard A. G. do, **Relatório técnico**: introdução à engenharia de software experimental. Rio de Janeiro: Programa de Engenharia de Sistemas e Computação, COPPE/UFRJ, 2002.

WANG, James Z.; TAYLOR, William. Concept Forest: A new ontology-assisted text document similarity measurement method. In: IEEE/WIC/ACM INTERNATIONAL CONFERENCE ON WEB INTELLIGENCE, 2007. **Proceedings...** 2007.

WENGER, Etienne. **Supporting communities of practice**: a survey of community-oriented technologies. 2001. Disponível em: <<http://www.ewenger.com/tech/index.htm>>. Acesso em: 1 jun. 2008.

YIMAM-SEID, Dawit; KOBASA, Alfred. Expert finding systems for organizations: Problem and Domain Analysis and the DEMOIR Approach. **Journal of Organizational Computing and Electronic Commerce**, v. 13, p. 1-24, 2004.

YONG-FENG, Shi; YAN-PING, Zhao. Comparison of text categorization algorithms. **WUJNS**, v. 9, n. 5, p 798-804, 2004.

APÊNDICE A – Documentos utilizados no estudo de viabilidade

Questionário 01 - Utilização de recursos de EAD

1. Por favor, relate a sua experiência com o ensino, tendo em vista as seguintes questões: quanto tempo de experiência em ensino, quando se formou, curso escolhido, titulação mais recente.
2. Em algum momento de sua formação docente (graduação, pós-graduação, mestrado, doutorado ou pós-doutorado), você teve contato com as tecnologias que favoreceram a construção do conhecimento em espaços virtuais, como por exemplo, ter participado de construção coletiva do conhecimento em fóruns de discussão?
3. Você tem experiência com EAD? Há quanto tempo trabalha neste segmento de ensino?
4. Em sua prática pedagógica (presencial ou à distância), você utiliza recursos didáticos, tais como: fóruns, chats, blogs, dentre outras ferramentas virtuais? Por quê?
5. Caso a pergunta acima seja positiva, responda:
 - a. Quais são os êxitos e dificuldades encontrados neste tipo interação, no contexto da relação ensino-aprendizagem?
 - b. Falando mais especificamente de listas de discussão, fale sobre estratégias de mediação utilizadas para mobilizar a participação dos estudantes.

Questionário 02 – Percepções dos professores participantes do estudo de viabilidade da ferramenta

1. Nome Completo:
2. Qual a sua formação?
3. Qual o nome da(s) Instituição(ões) de Ensino Superior na(s) qual(is) possui vínculo atualmente?
4. Que disciplina(s) leciona?
5. Por favor, relate a sua experiência em relação ao uso do FindYourHelp no momento de montar as hierarquias de categorias/termos referentes à sua disciplina. A interação com a ferramenta foi satisfatória? Você sugere alguma melhoria neste processo?
6. Sobre a identificação automática dos participantes especialistas. Na sua percepção, qual o grau de acerto da ferramenta?

 1 - Nenhum acerto 2 - Pouco acerto 3 - Acertou moderadamente

 4 - Acertou em sua maioria 5 - Acertou Totalmente
7. Em relação à análise individual feita sobre cada mensagem postada (categorização e descarte), a ferramenta acertou?

 1 - Nenhuma vez 2 - Poucas vezes 3 - Moderadamente

 4 - Muitas vezes 5 - Sempre
8. Em termos gerais, o quão confiável foi o resultado de todo o processo imposto pelo **FindYourHelp** em seu uso?

 1 - Não confiável 2 - Pouco confiável 3 - Confiável

 4 - Muito Confiável 5 - Totalmente confiável
9. De um modo geral, quais seriam as suas recomendações que possam trazer melhorias ao uso da ferramenta proposta?

APÊNDICE B – Código para importação do Openthesaurus

A seguir é exibido o código desenvolvido na linguagem Java para executar a importação dos termos do OptnThesaurus para a base de dados do FindYourHelp.

```

/*
 * OpenThesaurusMain.java
 *
 * Created on 17/11/2009, 23:23:25
 */
package gui;

import dal.ThesaurusDAO;
import java.awt.Cursor;
import java.io.File;
import java.io.FileNotFoundException;
import java.util.Hashtable;
import java.util.Scanner;
import java.util.logging.Level;
import java.util.logging.Logger;
import javax.swing.JFileChooser;

/**
 *
 * @author Marcos Santos
 */
public class OpenThesaurusMain extends javax.swing.JFrame {

    /** Creates new form OpenThesaurusMain */
    public OpenThesaurusMain() {
        initComponents();
    }

    @SuppressWarnings("unchecked")
    // <editor-fold defaultstate="collapsed" desc="Generated Code">
    private void initComponents() {

        lblFileUp = new javax.swing.JLabel();
        btnSearch = new javax.swing.JButton();
        txtFileUp = new javax.swing.JTextField();
        btnImport = new javax.swing.JButton();
        pbrStatusImport = new javax.swing.JProgressBar();
        lblStatus = new javax.swing.JLabel();
        setDefaultCloseOperation(javax.swing.WindowConstants.EXIT_ON_CLOSE);
        setTitle("OpenThesaurusPT Import");
        lblFileUp.setText("Escolha o arquivo:");
        btnSearch.setText("Procurar...");
        btnSearch.addActionListener(new java.awt.event.ActionListener() {
            public void actionPerformed(java.awt.event.ActionEvent evt) {
                btnSearchActionPerformed(evt);
            }
        });

        btnImport.setText("Importar dados do arquivo para o banco");
        btnImport.addActionListener(new java.awt.event.ActionListener() {
            public void actionPerformed(java.awt.event.ActionEvent evt) {
                btnImportActionPerformed(evt);
            }
        });

        javax.swing.GroupLayout layout = new
        javax.swing.GroupLayout(getContentPane());
        getContentPane().setLayout(layout);
        layout.setHorizontalGroup(

```

```

        layout.createParallelGroup(javax.swing.GroupLayout.Alignment.LEADING)
            .addGroup(javax.swing.GroupLayout.Alignment.TRAILING,
layout.createSequentialGroup()
            .addGap(25, 25, 25)

        .addGroup(layout.createParallelGroup(javax.swing.GroupLayout.Alignment.TRAILING)
            .addComponent(lblStatus,
javax.swing.GroupLayout.Alignment.LEADING, javax.swing.GroupLayout.DEFAULT_SIZE,
432, Short.MAX_VALUE)
            .addComponent(pbrStatusImport,
javax.swing.GroupLayout.Alignment.LEADING, javax.swing.GroupLayout.DEFAULT_SIZE,
432, Short.MAX_VALUE)
            .addComponent(btnImport, javax.swing.GroupLayout.DEFAULT_SIZE,
432, Short.MAX_VALUE)
            .addGroup(layout.createSequentialGroup()
                .addComponent(lblFileUp)
                .addGap(18, 18, 18)
                .addComponent(txtFileUp,
javax.swing.GroupLayout.DEFAULT_SIZE, 223, Short.MAX_VALUE)
                .addGap(18, 18, 18)
                .addComponent(btnSearch)))
            .addGap(32, 32, 32)
        );
        layout.setVerticalGroup(
            layout.createParallelGroup(javax.swing.GroupLayout.Alignment.LEADING)
                .addGroup(layout.createSequentialGroup()
                    .addGap(38, 38, 38)

        .addGroup(layout.createParallelGroup(javax.swing.GroupLayout.Alignment.LEADING)
            .addComponent(btnSearch)

        .addGroup(layout.createParallelGroup(javax.swing.GroupLayout.Alignment.BASELINE)
            .addComponent(lblFileUp)
            .addComponent(txtFileUp,
javax.swing.GroupLayout.PREFERRED_SIZE, javax.swing.GroupLayout.DEFAULT_SIZE,
javax.swing.GroupLayout.PREFERRED_SIZE)))
            .addGap(18, 18, 18)
            .addComponent(btnImport)
            .addGap(18, 18, 18)
            .addComponent(pbrStatusImport,
javax.swing.GroupLayout.PREFERRED_SIZE, javax.swing.GroupLayout.DEFAULT_SIZE,
javax.swing.GroupLayout.PREFERRED_SIZE)
            .addGap(18, 18, 18)
            .addComponent(lblStatus, javax.swing.GroupLayout.PREFERRED_SIZE,
23, javax.swing.GroupLayout.PREFERRED_SIZE)
            .addContainerGap(36, Short.MAX_VALUE))
        );
        pack();
    } // </editor-fold>

    private void btnSearchActionPerformed(java.awt.event.ActionEvent evt) {
        JFileChooser openFile = new JFileChooser();
        int op = openFile.showOpenDialog(this);
        if (op == JFileChooser.APPROVE_OPTION) {
            txtFileUp.setText(openFile.getSelectedFile().getAbsolutePath());
        }
    }

    private void btnImportActionPerformed(java.awt.event.ActionEvent evt) {
        importar(txtFileUp.getText());
    }

    public void importar(String filePath) {
        File f = new File(filePath);

        Cursor cursor = Cursor.getPredefinedCursor(Cursor.WAIT_CURSOR);
        setCursor(cursor);
    }

```

```

        lblStatus.setText("Importanto Arquivo...");

        Hashtable<String, SinonimosPalavra> thesaurus = new Hashtable<String,
SinonimosPalavra>();
        try {
            Scanner leitor = new Scanner(f);
            String linha;
            String[] palavras;
            // lê o arquivo linha a linha
            while(leitor.hasNextLine()){
                linha = leitor.nextLine();
                // Pula as linhas de comentários
                if((linha.length()>2) && !linha.startsWith("#") &&
!linha.substring(1,2).equals("#")) {
                    palavras = linha.split(";");
                    //Cria a instancia de SinonimoPalavra definindo a
                    //palavra base como seno a primeira da linha que foi
                    //lida, caso a palavra base já exista, adiciona as
                    //novas palavras ao objeto SinonimoPalavra existente
                    SinonimosPalavra sinP = new SinonimosPalavra();
                    if(thesaurus.containsKey(palavras[0].trim())){
                        sinP = (SinonimosPalavra) thesaurus.get(
palavras[0].trim());
                    }else{
                        sinP.setPalavraBase(palavras[0].trim());
                    }
                    for(int i=1; i<palavras.length; i++){
                        // Pula eventuais repetições de palavras base
                        if(!palavras[i].trim().equals(sinP.getPalavraBase())){
                            sinP.addSinonimo(palavras[i].trim());
                        }
                    }
                    thesaurus.put (palavras[0].trim(), sinP);
                }
            }
            new ThesaurusDAO().inserirBatch(thesaurus);
            lblStatus.setText("Arquivo importado com sucesso!");
        } catch (FileNotFoundException ex) {
            Logger.getLogger(OpenThesaurusMain.class.getName()).log(Level.SEVERE,
null, ex);
            lblStatus.setText("Ocorreu um erro importante o Arquivo!");
            "+ex.getMessage());
        } catch(Exception ex){
            Logger.getLogger(OpenThesaurusMain.class.getName()).log(Level.SEVERE,
null, ex);
            lblStatus.setText("Ocorreu um erro importante o Arquivo!");
            "+ex.getMessage());
        } finally{
            cursor = Cursor.getDefaultCursor();
            setCursor( cursor );
        }
    }
    /**
     * @param args the command line arguments
     */
    public static void main(String args[]) {
        java.awt.EventQueue.invokeLater(new Runnable() {
            public void run() {
                new OpenThesaurusMain().setVisible(true);
            }
        });
    }

    private javax.swing.JButton btnImport;
    private javax.swing.JButton btnSearch;
    private javax.swing.JLabel lblFileUp;
    private javax.swing.JLabel lblStatus;
    private javax.swing.JProgressBar pbrStatusImport;

```

```

    private javax.swing.JTextField txtFileUp;
}

package gui;

import java.util.ArrayList;

/**
 *
 * @author Marcos
 */
public class SinonimosPalavra {
    private String palavraBase;
    private ArrayList<String> sinonimos = new ArrayList<String>();

    /**
     * @return the palavraBase
     */
    public String getPalavraBase() {
        return palavraBase;
    }

    /**
     * @param palavraBase the palavraBase to set
     */
    public void setPalavraBase(String palavraBase) {
        this.palavraBase = palavraBase;
    }

    /**
     * @return the sinonimos
     */
    public ArrayList<String> getSinonimos() {
        return sinonimos;
    }

    /**
     * @param sinonimos the sinonimos to set
     */
    public void setSinonimos(ArrayList<String> sinonimos) {
        this.sinonimos = sinonimos;
    }

    public void addSinonimo(String sin){
        if(!sinonimos.contains(sin)){
            sinonimos.add(sin);
        }
    }
}

package dal;

/**
 *
 * @author Marcos
 */
public class ConstantesDB {
    public static final String URL = "jdbc:mysql://127.0.0.1:3306/moodle19";
    public static final String DRV = "com.mysql.jdbc.Driver";
    public static final String USER = "root";
    public static final String PWD = "*****";
}

```



```

package dal;

import gui.SinonimosPalavra;
import java.sql.Connection;
import java.sql.DriverManager;
import java.sql.PreparedStatement;
import java.sql.ResultSet;
import java.sql.SQLException;
import java.util.Hashtable;

/**
 *
 * @author Marcos
 */
public class ThesaurusDAO {
    public void inserirBatch(Hashtable<String, SinonimosPalavra> palavras){

        Connection con = null;

        try {
            Class.forName(ConstantsDB.DRV).newInstance();
            con = DriverManager.getConnection(ConstantsDB.URL, ConstantsDB.USER,
ConstantsDB.PWD);

            for(SinonimosPalavra sp : palavras.values()){

                // Inseire a palavra base
                PreparedStatement ps = con.prepareStatement("INSERT INTO
mdl_fyh_thesaurus_bw (baseword) VALUES (?)");
                ps.setString(1, sp.getPalavraBase());
                ps.executeUpdate();

                // Obtem o id inserido
                ps = con.prepareStatement("SELECT id FROM mdl_fyh_thesaurus_bw WHERE
baseword = ?");
                ps.setString(1, sp.getPalavraBase());
                ResultSet rs = ps.executeQuery();

                int idWB = 0;
                while(rs.next()){
                    idWB = rs.getInt("id");
                    break;
                }

                // Inseire os sinônimos da Palavra Base
                for(String syn : sp.getSinonimos()){
                    ps = con.prepareStatement("INSERT INTO mdl_fyh_thesaurus_syn
(id_baseword, syn) VALUES (?, ?)");
                    ps.setInt(1, idWB);
                    ps.setString(2, syn);
                    ps.executeUpdate();
                }
            }
        } catch(Exception e) {
            System.err.println("Exception: " + e.getMessage());
        } finally {
            try {
                if(con != null)
                    con.close();
            } catch(SQLException e) {}
        }
    }
}

```

APÊNDICE C – Sinais Gráficos Removidos

A seguir são exibidos os sinais gráficos removidos do texto das mensagens durante a fase de pré-processamento do algoritmo de mineração de texto. Há um tratamento especial para o ponto e o hífen para que as palavras compostas e e-mails não sejam alterados.

, . : ? ! ; () { } \ / [] = +
- # \$ * & % 0 1 2 3 4 5 6 7 8 9

ANEXO A – Respostas aos questionários do estudo

Respostas do Questionário 01 - Utilização de recursos de EAD

Professor(a): A. J. A. C.

6. **Por favor, relate a sua experiência com o ensino, tendo em vista as seguintes questões: quanto tempo de experiência em ensino, quando se formou, curso escolhido, titulação mais recente.**

5 Anos de Ensino, formado em 2004, Especialização em Administração de Banco de Dados

7. **Em algum momento de sua formação docente (graduação, pós-graduação, mestrado, doutorado ou pós-doutorado), você teve contato com as tecnologias que favoreceram a construção do conhecimento em espaços virtuais, como por exemplo, ter participado de construção coletiva do conhecimento em fóruns de discussão?**

SIM

8. **Você tem experiência com EAD? Há quanto tempo trabalha neste segmento de ensino?**

3 Anos

9. **Em sua prática pedagógica (presencial ou à distância), você utiliza recursos didáticos, tais como: fóruns, chats, blogs, dentre outras ferramentas virtuais? Por quê?**

SIM. Porque é possível ampliar os conteúdos de sala de aula, despertar nos alunos a capacidade de investigação e aprimorar o desenvolvimento de tarefas remotamente.

10. **Caso a pergunta acima seja positiva, responda:**

- a. **Quais são os êxitos e dificuldades encontrados neste tipo interação, no contexto da relação ensino-aprendizagem?**

Quebra de Paradigmas, gestão de informações analiticamente, rapidez na recuperação da informação histórica já discutida em anos anteriores.

- b. **Falando mais especificamente de listas de discussão, fale sobre estratégias de mediação utilizadas para mobilizar a participação dos estudantes.**

O processo de ensino aprendizagem é desenvolvido através de tópicos e questionamentos, que buscam despertar a reflexão com indagações técnicas. Onde o senso crítico do aluno será avaliado.

Professor(a): E. R.

1. **Por favor, relate a sua experiência com o ensino, tendo em vista as seguintes questões: quanto tempo de experiência em ensino, quando se formou, curso escolhido, titulação mais recente.**

Experiência: 10 anos

Formação: 1992

Curso Matemática

Titulação recente: Mestre em Modelagem Computacional

2. **Em algum momento de sua formação docente (graduação, pós-graduação, mestrado, doutorado ou pós-doutorado), você teve contato com as tecnologias que favoreceram a construção do conhecimento em espaços virtuais, como por exemplo, ter participado de construção coletiva do conhecimento em fóruns de discussão?**

Sim, utilizando o Moodle e o Teleduc

3. **Você tem experiência com EAD? Há quanto tempo trabalha neste segmento de ensino?**

Experiência apenas como aluno

4. **Em sua prática pedagógica (presencial ou à distância), você utiliza recursos didáticos, tais como: fóruns, chats, blogs, dentre outras ferramentas virtuais? Por quê?**

Sim, facilita a comunicação e rompe com as barreiras geográficas e temporais

5. **Caso a pergunta acima seja positiva, responda:**

- a) **Quais são os êxitos e dificuldades encontrados neste tipo interação, no contexto da relação ensino-aprendizagem?**

São muitas as vantagens, que vai desde a não perda de tempo com engarrafamento, a disponibilidade de fazer as atividades em momentos mais convenientes, a necessidade de maior planejamento, disciplina, responsabilidade e assiduidade por parte do aluno entre outras.

As dificuldades dizem respeito principalmente a irresponsabilidade, mas aí já depende muito mais do aluno e algumas vezes a tecnologia disponível em cada caso.

- b) **Falando mais especificamente de listas de discussão, fale sobre estratégias de mediação utilizadas para mobilizar a participação dos estudantes.**

Controlar número de acesso, tempo de acessos, participações em chats e fóruns, qualidade das participações e número de acerto em questões.

Professor(a): E. S. S.

- 1. Por favor, relate a sua experiência com o ensino, tendo em vista as seguintes questões: quanto tempo de experiência em ensino, quando se formou, curso escolhido, titulação mais recente.**

Quanto tempo de experiência em ensino: 5 anos

Quando se formou: 06/2000

Curso escolhido: Processamento de Dados

Titulação mais recente: Especialização

- 2. Em algum momento de sua formação docente (graduação, pós-graduação, mestrado, doutorado ou pós-doutorado), você teve contato com as tecnologias que favoreceram a construção do conhecimento em espaços virtuais, como por exemplo, ter participado de construção coletiva do conhecimento em fóruns de discussão?**

Na formação não. Tive na atuação como docente (Ambiente Moodle).

- 3. Você tem experiência com EAD? Há quanto tempo trabalha neste segmento de ensino?**

Não.

- 4. Em sua prática pedagógica (presencial ou à distância), você utiliza recursos didáticos, tais como: fóruns, chats, blogs, dentre outras ferramentas virtuais? Por quê?**

Apenas grupos de e-mail e/ou listas de discussão. Falta de infra-estrutura, institucional, montada e disponível para uso. Indisponibilidade, pessoal, para montagem de infra-estrutura com recursos próprios.

- 5. Caso a pergunta acima seja positiva, responda:**

- a) Quais são os êxitos e dificuldades encontrados neste tipo interação, no contexto da relação ensino-aprendizagem?**

Êxitos: Maior participação da turma na exposição de opiniões e experiências que vem a enriquecer o curso.

Dificuldades: Motivá-los a participar. Realizar o pontapé inicial. Alguns estudantes também postam mensagens apenas para contar mais uma participação, para o professor pensar que estão interagindo muito,

porém o que se percebe é que as mensagens vêm sem muito conteúdo. Isso dificulta o trabalho de avaliação do professor, pois em meio a muitas mensagens, poucas terminam realmente discutindo algo interessante.

b) Falando mais especificamente de listas de discussão, fale sobre estratégias de mediação utilizadas para mobilizar a participação dos estudantes.

As listas foram abertas aos participantes e sem regras de mediação. Quanto à motivação, deve-se envolver os alunos com assuntos que mexam com eles e os levem a expor as suas idéias num ambiente colaborativo com o objetivo de participação na evolução de conhecimento. Eles devem ser convidados a participar das discussões e nós, como docentes e mediadores, devemos encorajá-los a contribuir com as suas experiências mostrando que são importantes e que agregam valor.

Professor(a): H. A.

1. Por favor, relate a sua experiência com o ensino, tendo em vista as seguintes questões: quanto tempo de experiência em ensino, quando se formou, curso escolhido, titulação mais recente.

Tempo de experiência: 4 anos
 Ano de formação 2003
 Ciência da computação
 Mestre

2. Em algum momento de sua formação docente (graduação, pós-graduação, mestrado, doutorado ou pós-doutorado), você teve contato com as tecnologias que favoreceram a construção do conhecimento em espaços virtuais, como por exemplo, ter participado de construção coletiva do conhecimento em fóruns de discussão?

Sim.

3. Você tem experiência com EAD? Há quanto tempo trabalha neste segmento de ensino?

Um pouco. Trabalhei em um centro universitário montando o ambiente EAD. Este trabalho durou 8 meses

4. Em sua prática pedagógica (presencial ou à distância), você utiliza recursos didáticos, tais como: fóruns, chats, blogs, dentre outras ferramentas virtuais? Por quê?

Sim. Utilizo fóruns e meu próprio site. Isto permite aumentar a dinâmica da disciplina e incrementa o compartilhamento de informações entre os alunos / professor.

5. Caso a pergunta acima seja positiva, responda:

a) Quais são os êxitos e dificuldades encontrados neste tipo interação, no contexto da relação ensino-aprendizagem?

Êxitos: Compartilhamento de informação e conhecimento. Com as tecnologias virtuais a disciplina não se restringe apenas na sala de aula.

Dificuldades: em convencer ao aluno que o conteúdo de internet não deve ser copiado facilmente. Orientar ao aluno qual conteúdo virtual é adequado ou não para o uso.

b) Falando mais especificamente de listas de discussão, fale sobre estratégias de mediação utilizadas para mobilizar a participação dos estudantes.

Colocar desafios e textos para que os alunos reflitam e opinem virtualmente.

Professor(a): M.P.

1. Por favor, relate a sua experiência com o ensino, tendo em vista as seguintes questões: quanto tempo de experiência em ensino, quando se formou, curso escolhido, titulação mais recente.

Experiência: 11 anos, sendo 7 em ensino superior.

Formação: Ciências Contábeis em 1992

Titulação mais recente: Pós-graduação em Sistemas de Informação com ênfase em banco de dados – Faculdade Ruy Barbosa - 2000

2. Em algum momento de sua formação docente (graduação, pós-graduação, mestrado, doutorado ou pós-doutorado), você teve contato com as tecnologias que favoreceram a construção do conhecimento em espaços virtuais, como por exemplo, ter participado de construção coletiva do conhecimento em fóruns de discussão?

Em minha formação tive contato em cursos de extensão.

3. Você tem experiência com EAD? Há quanto tempo trabalha neste segmento de ensino?

Trabalho com EAD desde 2007.

- 4. Em sua prática pedagógica (presencial ou à distância), você utiliza recursos didáticos, tais como: fóruns, chats, blogs, dentre outras ferramentas virtuais? Por quê?**

Sim. Pois facilitam a construção do conhecimento e o acompanhamento do curso.

- 5. Caso a pergunta acima seja positiva, responda:**

- a) Quais são os êxitos e dificuldades encontrados neste tipo interação, no contexto da relação ensino-aprendizagem?**

Quando utilizados isoladamente contribuem pouco em relação ao processo, porém, quando utilizados em conjunto e como complemento e apoio às aulas presenciais trazem grandes benefícios.

- b) Falando mais especificamente de listas de discussão, fale sobre estratégias de mediação utilizadas para mobilizar a participação dos estudantes.**

Deixar as listas para utilização esporádica e opcional mobiliza muito pouco os alunos. O uso mais eficaz ocorre sempre quando sua utilização faça parte do próprio andamento e acompanhamento do curso, inclusive em relação a avaliação.

Professor(a): M.V.

- 1. Por favor, relate a sua experiência com o ensino, tendo em vista as seguintes questões: quanto tempo de experiência em ensino, quando se formou, curso escolhido, titulação mais recente.**

Quanto tempo de experiência em ensino: 3 anos
 Quando se formou Graduação: 2006; Pós-graduação: 2008
 Curso Escolhido: Pós-graduação em Sistemas Distribuídos
 Titulação mais recente: Pós-graduado (Lato Sensu)

- 2. Em algum momento de sua formação docente (graduação, pós-graduação, mestrado, doutorado ou pós-doutorado), você teve contato com as tecnologias que favoreceram a construção do conhecimento em espaços virtuais, como por exemplo, ter participado de construção coletiva do conhecimento em fóruns de discussão?**

Sim. Programa de Formação Docente fornecido pela Estácio.

3. Você tem experiência com EAD? Há quanto tempo trabalha neste segmento de ensino?

Sim. Atuei como professor conteudista e tutor em curso de extensão pela Unifacs durante 06 meses.

4. Em sua prática pedagógica (presencial ou à distância), você utiliza recursos didáticos, tais como: fóruns, chats, blogs, dentre outras ferramentas virtuais? Por quê?

Sim. Utilizo-os como instrumentos que favorecem a interatividade ultrapassando as barreiras físicas de uma sala de aula.

5. Caso a pergunta acima seja positiva, responda:

a) Quais são os êxitos e dificuldades encontrados neste tipo interação, no contexto da relação ensino-aprendizagem?

Êxitos: “democratização” da educação. Dificuldades: resistência à nova modalidade de ensino-aprendizagem tanto pelos alunos quanto pelos coordenadores, professores e tutores.

b) Falando mais especificamente de listas de discussão, fale sobre estratégias de mediação utilizadas para mobilizar a participação dos estudantes.

Através de provocações, incentivando a contribuição e participação do coletivo e de cada aluno em particular.

Professor(a): O. B. N.

1. Por favor, relate a sua experiência com o ensino, tendo em vista as seguintes questões: quanto tempo de experiência em ensino, quando se formou, curso escolhido, titulação mais recente.

Resposta: sou Mestre em Informática desde 1999, Bacharel em Ciência da Computação desde 1995 e Estudante de Doutorado em Ciência da Computação desde 2007. Leciono em instituições de ensino superior há cerca de 10 anos e em instituições de ensino profissionalizante/técnico por 5 anos.

2. Em algum momento de sua formação docente (graduação, pós-graduação, mestrado, doutorado ou pós-doutorado), você teve contato com as tecnologias que favoreceram a construção do conhecimento em espaços virtuais, como por exemplo, ter participado de construção coletiva do conhecimento em fóruns de discussão?

Não.

3. Você tem experiência com EAD? Há quanto tempo trabalha neste segmento de ensino?

Sim, trabalho com o ambiente Moodle há mais de cinco anos.

4. Em sua prática pedagógica (presencial ou à distância), você utiliza recursos didáticos, tais como: fóruns, chats, blogs, dentre outras ferramentas virtuais? Por quê?

Sim. Utilizo os recursos didáticos disponibilizados no Ambiente de Virtual de Aprendizagem (AVA) Moodle. Os recursos são usados para que a interação com o aluno seja a maior possível.

5. Caso a pergunta acima seja positiva, responda:

a) Quais são os êxitos e dificuldades encontrados neste tipo interação, no contexto da relação ensino-aprendizagem?

Interessante citar, que mesmo alunos tímidos em aulas presenciais tendem a ser mais extrovertidos em AVA. Além disso, o aluno tende a elaborar um horário de estudos adequado às necessidades dele. O que, por sua vez, também pode ser uma desvantagem quando o aluno é desorganizado por natureza. Os alunos ainda acostumados ao ambiente de ensino tradicional tendem a ter alguma dificuldade no início do processo, entretanto, graças à inserção da Internet na vida de quase todos os alunos, logo estes suplantam as dificuldades.

b) Falando mais especificamente de listas de discussão, fale sobre estratégias de mediação utilizadas para mobilizar a participação dos estudantes.

O meio mais efetivo de garantir a participação dos alunos nas listas de discussão foi através da implantação de uma regra que garante pontos extras para os alunos com participações mais úteis. Isso se deve ao fato do nosso aluno ainda crer que nota é o item mais importante no processo ensino-aprendizagem.

Professor(a): T. D.

1. Por favor, relate a sua experiência com o ensino, tendo em vista as seguintes questões: quanto tempo de experiência em ensino, quando se formou, curso escolhido, titulação mais recente.

Quanto tempo de experiência em ensino: 10 anos (cursos de informática - 3, ensino Fundamental, Médio e Profissionalizante – 6, Ensino Superior - 3)
 Quando se formou: 2005
 Curso escolhido: Bacharelado em Sistemas de Informação
 Titulação mais recente: Especialista em Sistemas de Informação – cursando o mestrado

2. Em algum momento de sua formação docente (graduação, pós-graduação, mestrado, doutorado ou pós-doutorado), você teve contato com as tecnologias que favoreceram a construção do conhecimento em espaços virtuais, como por exemplo, ter participado de construção coletiva do conhecimento em fóruns de discussão?

Sim. Lista de discussão na graduação com professores e colegas de classe; Lista de discussão na especialização com colegas de classe; Lista de discussão e Moodle no mestrado com professores e colegas de classe.

3. Você tem experiência com EAD? Há quanto tempo trabalha neste segmento de ensino?

Sim. Há quatro anos. Dois cursos preparatórios (UNEB e UFBA) – 2006-2007; tema da monografia da especialização- 2007-2008; ministrei disciplina à distância pela FIB (Moodle) e Estácio (SIA) – 2008-2009; cursos de formação continuada de professores da Estácio/FIB – 2009-Atual.

4. Em sua prática pedagógica (presencial ou à distância), você utiliza recursos didáticos, tais como: fóruns, chats, blogs, dentre outras ferramentas virtuais? Por quê?

Sim. À distância: fóruns, lista de discussão, wiki, biblioteca, glossário, upload de arquivos, agenda/calendário, links. Obs.: como foi o tema da monografia da especialização tentei usar todos os recursos do Moodle.
 Presencial: lista de discussões e links.

5. Caso a pergunta acima seja positiva, responda:

a) Quais são os êxitos e dificuldades encontrados neste tipo interação, no contexto da relação ensino-aprendizagem?

Êxitos: fácil acesso à informação; a informação é disponibilizada de forma variada e mais atrativa; o aluno reflete melhor sobre o processo da aprendizagem, torna-se mais crítico e criativo; o aluno é ator do processo de ensino-aprendizagem; como são pessoas que estão distantes, sentem-se à vontade para ter dúvidas; conhecem novas culturas, caso seja um curso distante fisicamente.

Dificuldades: muitos alunos copiam trechos da internet e postam como de sua autoria sem refletir sobre o seu conteúdo nem direitos autorais; dificuldade de acesso (banda e configuração do hardware);

conscientizar o aluno de que ele é responsável pela construção do seu conhecimento.

b) Falando mais especificamente de listas de discussão, fale sobre estratégias de mediação utilizadas para mobilizar a participação dos estudantes.

Trabalho com pontuação qualitativa e quantitativa: qualidade da informação postada, número de acesso (informação postada, leitura, correção), tempo utilizado para postar e ler a informação, perfil que o aluno cria ao longo da disciplina pode ser utilizado para identificar o aluno; provocações através de questionamento sobre a postagem do aluno solicitando sua resposta e parecer do grupo, no final de cada unidade o aluno copia as postagens e elabora uma dissertação, caso use as dos colegas deve citar o autor.

Respostas do Questionário 02 - Percepções dos professores participantes do estudo de viabilidade da ferramenta

1. Nome Completo:

E. S.

2. Qual a sua formação?

Especialista e Sistemas de Informação com ênfase em Componentes Distribuídos e Web

3. Qual o nome da(s) Instituição(ões) de Ensino Superior na(s) qual(is) possui vínculo atualmente?

Centro Universitário da Bahia – FIB e Faculdade Ruy Barbosa - FRB

4. Que disciplina(s) leciona?

Projeto Avançado de Sistemas I e II

5. Por favor, relate a sua experiência em relação ao uso do FindYourHelp no momento de montar as hierarquias de categorias/termos referentes à sua disciplina. A interação com a ferramenta foi satisfatória? Você sugere alguma melhoria neste processo?

A montagem da hierarquia de termos do FindYourHelp é simples, porém precisa melhorar no quesito de alteração das categorias, por exemplo: um termo pertencente a uma categoria não pode ser movido ou copiado para outra.

6. Sobre a identificação automática dos participantes especialistas. Na sua percepção, qual o grau de acerto da ferramenta?

() 1 - Nenhum acerto () 2 - Pouco acerto () 3 - Acertou moderadamente

4 - Acertou em sua maioria 5 - Acertou Totalmente

7. Em relação à análise individual feita sobre cada mensagem postada (categorização e descarte), a ferramenta acertou?

1 - Nenhuma vez 2 - Poucas vezes 3 - Moderadamente

4 - Muitas vezes 5 - Sempre

8. Em termos gerais, o quão confiável foi o resultado de todo o processo imposto pelo FindYourHelp em seu uso?

1 - Não confiável 2 - Pouco confiável 3 - Confiável

4 - Muito Confiável 5 - Totalmente confiável

9. De um modo geral, quais seriam as suas recomendações que possam trazer melhorias ao uso da ferramenta proposta?

Como já disse, a questão da alteração das categorias poderia ser uma alteração interessante. Outra coisa importante seria o uso de funcionalidades mais ricas de interface com o usuário, como arrastar e soltar termos de uma categoria para outra.

1. Nome Completo:

H. A.

2. Qual a sua formação?

Mestre em Sistemas e Computação

3. Qual o nome da(s) Instituição(ões) de Ensino Superior na(s) qual(is) possui vínculo atualmente?

Centro Universitário da Bahia – FIB

4. Que disciplina(s) leciona?

Linguagem de Programação Orientada a Objetos I e II

5. Por favor, relate a sua experiência em relação ao uso do FindYourHelp no momento de montar as hierarquias de categorias/termos referentes à sua disciplina. A interação com a ferramenta foi satisfatória? Você sugere alguma melhoria neste processo?

A interação com o usuário está fácil, porém percebi que não podemos definir termos compostos por mais de uma palavra, como por exemplo: “classe abstrata”, este tipo de termo poderia ser mais útil na identificação de um assunto relacionado à minha disciplina sendo comparados com um só do que separadamente.

6. Sobre a identificação automática dos participantes especialistas. Na sua percepção, qual o grau de acerto da ferramenta?

1 - Nenhum acerto 2 - Pouco acerto 3 - Acertou moderadamente

4 - Acertou em sua maioria 5 - Acertou Totalmente

7. Em relação à análise individual feita sobre cada mensagem postada (categorização e descarte), a ferramenta acertou?

1 - Nenhuma vez 2 - Poucas vezes 3 - Moderadamente

4 - Muitas vezes 5 - Sempre

8. Em termos gerais, o quão confiável foi o resultado de todo o processo imposto pelo FindYourHelp em seu uso?

1 - Não confiável 2 - Pouco confiável 3 - Confiável

4 - Muito Confiável 5 - Totalmente confiável

9. De um modo geral, quais seriam as suas recomendações que possam trazer melhorias ao uso da ferramenta proposta?

A ferramenta poderia sugerir o contato de uma pessoa no momento em que a mensagem está sendo digitada, através de algum recurso como AJAX, por exemplo. Talvez isso deixe o algoritmo mais lento, porém, se implementado, pode trazer um dinamismo maior para o usuário.

1. Nome Completo:

J. F.

2. Qual a sua formação?

Mestre em História da Educação

3. Qual o nome da(s) Instituição(ões) de Ensino Superior na(s) qual(is) possui vínculo atualmente?

Centro Universitário da Bahia – FIB e Faculdade Ruy Barbosa - FRB

4. Que disciplina(s) leciona?

Metodologia da Pesquisa I e II

5. Por favor, relate a sua experiência em relação ao uso do FindYourHelp no momento de montar as hierarquias de categorias/termos referentes à sua disciplina. A interação com a ferramenta foi satisfatória? Você sugere alguma melhoria neste processo?

Precisei da ajuda do autor deste trabalho na elaboração da hierarquia por ser uma nova forma de perceber como está estruturada a disciplina a ser analisada, quanto ao modo de interagir com o sistema, achei de fácil compreensão. Entretanto, em minha opinião, a ferramenta deveria permitir a digitação de termos compostos de várias palavras, quando estas possuem um sentido maior para definir o conceito abordado na disciplina estando vinculadas como se fora um único termo.

6. Sobre a identificação automática dos participantes especialistas. Na sua percepção, qual o grau de acerto da ferramenta?

1 - Nenhum acerto 2 - Pouco acerto 3 - Acertou moderadamente

4 - Acertou em sua maioria 5 - Acertou Totalmente

7. Em relação à análise individual feita sobre cada mensagem postada (categorização e descarte), a ferramenta acertou?

1 - Nenhuma vez 2 - Poucas vezes 3 - Moderadamente

4 - Muitas vezes 5 - Sempre

8. Em termos gerais, o quão confiável foi o resultado de todo o processo imposto pelo FindYourHelp em seu uso?

1 - Não confiável 2 - Pouco confiável 3 - Confiável

4 - Muito Confiável 5 - Totalmente confiável

9. De um modo geral, quais seriam as suas recomendações que possam trazer melhorias ao uso da ferramenta proposta?

Penso que faltou uma prática maior com o sistema para tecer um comentário mais elaborado neste sentido, porém entendo que o FindYourHelp poderia permitir que os estudantes e professores consultassem os especialistas digitando o nome do assunto que desejam, pois procurar na estrutura de árvore atual pode se tornar complicado, caso a hierarquia definida seja muito grande. Outra sugestão é criar algum tipo de controle dentro do sistema para os casos em que um estudante posta em nome de um grupo inteiro, pois a pontuação, neste caso, fica atrelada a apenas um participante.