

Aplicando técnicas de Recuperação de Informações para facilitar a interação em ambientes cooperativos: uma abordagem multiagentes

Renata Silva Souza Guizzardi
Crediné Silva de Menezes
renataguizzardi@hotmail.com
credine@inf.ufes.br

Universidade Federal do Espírito Santo – UFES
Centro Tecnológico – Mestrado em Informática
Av. Fernando Ferrari, s/n
Campus Universitário de Goiabeiras, CT-VII
CEP: 29060-900, Vitória/ES, Tel. : (27) 335-2689

RESUMO

Cada vez mais, as pesquisas acadêmicas vêm apontando os benefícios da adoção do aprendizado colaborativo e, por isso, a área de Informática Educativa tem concentrado esforços no desenvolvimento de ambientes adequados a essa prática. Hoje, existem diversas pesquisas em Recuperação de Informações (RI) que apontam caminhos interessantes para a construção de ferramentas que apoiam operações importantes no contexto educacional. Este trabalho propõe um sistema multiagentes que utiliza técnicas de RI para apoiar a interação social em ambientes cooperativos de aprendizagem.

Palavras-chave: recuperação de informações, sistema multiagentes, interação, ambientes cooperativos de aprendizagem

1. Introdução

Cada vez mais, as pesquisas acadêmicas vêm apontando os benefícios da adoção de práticas educacionais ligadas ao aprendizado colaborativo que, segundo Jaffe (apud Souto et al., 1999) envolve três princípios importantes: interação, mediação e participação ativa do aluno. É importante lembrar, todavia, que o sucesso da aplicação de métodos colaborativos depende da criação de ambientes adequados para a cooperação, sejam eles presenciais ou não (Graves, 1994).

Os estudos na área de Informática Educativa têm se concentrado no desenvolvimento de ambientes adequados às novas práticas educacionais, que são conhecidos como Ambientes Cooperativos de Aprendizagem. Para que esses ambientes obtenham sucesso, é preciso que eles forneçam mais do que apenas ferramentas de comunicação, como email, fóruns de discussão e programas de chat. É necessário prover aos aprendizes recursos que possibilitem a organização e a manipulação das mensagens recebidas, de modo que eles possam encontrar conteúdos específicos e redigir mensagens de forma a expressar melhor suas idéias.

Hoje, existem, na Ciência da Computação, diversas pesquisas na área de Recuperação de Informações (RI) que podem apontar caminhos interessantes para a construção de ferramentas que dêem suporte às operações apresentadas. A área de

estudos em RI está relacionada à representação, ao armazenamento, à organização e ao acesso a itens de informação (Salton & McGill, 1983).

Este trabalho teve início com a análise de mensagens enviadas aos fóruns de discussão de um curso semi-presencial realizado pela Universidade Federal do Espírito Santo (UFES), entre os anos de 1999 e 2000. Os resultados dessa análise foram apresentados em (Souza & Menezes, 2000a). A partir dessa análise, concluiu-se alguns requisitos que levaram à proposta de um sistema para apoio à interação social em ambientes cooperativos de aprendizagem (Souza & Menezes, 2000b). Nesta oportunidade, pretende-se ampliar a proposta anterior, propondo um ambiente que apresente funções que dêem suporte não apenas à comunicação, mas também à organização do material enviado ao curso, à socialização dos participantes, entre outras facilidades.

Este artigo está organizado da seguinte forma: a seção 2 comenta sobre a necessidade de prover suporte à RI em ambientes de aprendizagem; a seção 3 apresenta o ambiente proposto; a seção 4 fala sobre o desenvolvimento de um protótipo desse ambiente e a seção 5 traz as considerações finais do trabalho.

2. Necessidades de RI em Ambientes de Aprendizagem

A partir do surgimento da Web, a Internet passou a ser um grande repositório de informações que pode ser explorado por cientistas, educadores, estudantes e outros usuários, como fonte alternativa aos tradicionais modelos de obtenção de informação, como, por exemplo, as bibliotecas.

Pode-se relacionar algumas vantagens do uso da Internet para a obtenção de informação científica, quando comparada aos métodos tradicionais (Blattmann et al., 2001): a facilidade e a rapidez no acesso à informação; a visão de outras realidades virtuais; o aumento da comunicação interpessoal; a maior interação e integração com outros alunos e professores; os ganhos no trabalho cooperativo, com a promoção de um maior intercâmbio; a melhoria da comunicação escrita; a produção de materiais, como a construção de documentos hipertexto; o aprendizado em diferentes conteúdos, visando um trabalho interdisciplinar.

Todos esses aspectos falam do grande potencial da Internet na recuperação de informação, porém é preciso investir em cada um deles para que as vantagens sejam efetivamente obtidas. Isso significa, de um lado, investir tempo para mudar as práticas educacionais e culturais para buscar novos caminhos que privilegiem a interação, a cooperação e o trabalho interdisciplinar, e também, significa criar novos sistemas e ferramentas baseadas no uso da Internet, que dêem suporte a todas essas atividades, para aumentar os ganhos e facilitar o uso da Rede.

Uma das principais preocupações com o uso da Internet na aprendizagem é a possibilidade de que os aprendizes se percam entre tantas opções e caminhos traçados pelo hipertexto. As ferramentas de busca hoje disponíveis na Internet estão longe de conseguir o tão almejado objetivo de recuperar documentos relevantes, a partir da necessidade do usuário. Entre as dificuldades enfrentadas, está o fato de que é impossível que uma dessas ferramentas cubra todo o conteúdo da Rede, assim, cada uma tem seu próprio índice, com um número limitado de documentos (Moreira, 1998). Além disso, essas ferramentas também são limitadas no que diz respeito à ordenação dos documentos encontrados de acordo com a preferência do usuário (ranqueamento) e, em geral, não apresentam uma interface que favoreça o usuário a expressar sua necessidade de informação (Baeza-Yates & Ribeiro-Neto, 1999).

É importante que os educadores e os profissionais ligados à aplicação da informática na educação aprendam um pouco sobre essa área de conhecimento para

que possam contribuir para a especificação de ferramentas mais apropriadas para o contexto educacional. Na verdade, já existem algumas iniciativas nessa direção, como os sistemas apresentados em (Lopes & Menezes, 2000) e (Castro Jr. & Silva, 2000).

É comum pensar em RI apenas no contexto de textos e documentos inteiros, entretanto, ferramentas de comunicação também precisam de recursos de RI, para recuperar mensagens que tratem de assuntos específicos. Este trabalho propõe um ambiente que utiliza técnicas de RI para dar suporte à interação entre aprendizes em Ambientes Cooperativos de Aprendizagem

3. Proposta de um Ambiente

A proposta do ambiente partiu da análise da ferramenta oferecida pelo WebCT (www.webct.com) para criação de fóruns de discussão. Esse foi o sistema de autoria utilizado no desenvolvimento do Curso de Especialização em Informática Educativa, oferecido pelo Departamento de Informática da UFES aos professores da Rede Pública Estadual de Educação do ES, entre os anos de 1999 e 2000. Essa foi uma das iniciativas associadas ao Programa Nacional de Informática na Educação (PROINFO), que objetiva a introdução da tecnologia da informação nas escolas públicas de todo o país como ferramenta de apoio ao processo ensino-aprendizagem. O curso formou 20 professores multiplicadores, que dividirão o conteúdo aprendido com outros professores de suas escolas. Maiores informações sobre a metodologia utilizada no curso e o ambiente cooperativo em que ele se realizou podem ser encontradas em (Souza & Menezes, 2000a).

3.1. Análise de uma ferramenta existente

Os fóruns de discussão que conhecemos hoje apresentam-se sob a forma de páginas Web, que exibem seqüências de títulos (assuntos) das mensagens enviadas por seus usuários, em geral, organizando-as em ordem cronológica em relação a data de envio.

Os principais problemas encontrados na ferramenta analisada podem ser sumarizados nos seguintes tópicos:

- O ambiente permite a criação de vários fóruns com temas diferentes e, depois, os próprios participantes decidem aonde querem postar as mensagens. A ferramenta, portanto, não oferece uma maneira de controlar em que fórum a mensagem está sendo postada, assim, é comum encontrar mensagens em um fórum que não dizem respeito ao assunto nele tratado;
- O ambiente não traz recursos para ajudar os participantes do curso a encontrarem uma resposta para suas dúvidas, assim, com o aumento do número de mensagens, é muito difícil encontrar o conteúdo requerido, o que leva os participantes a simplesmente repetirem uma mesma pergunta diversas vezes.
- A única maneira de identificar uma mensagem é pelo seu campo assunto e, como o autor da mensagem nem sempre indica claramente no assunto de sua mensagem o que ele escreveu em seu conteúdo, essa maneira é bastante ineficaz.
- A única função de configuração disponibilizada no ambiente é a que permite que o usuário escolha se quer visualizar todas as mensagens ou apenas as que ele ainda não leu.

Um ponto forte encontrado no ambiente foi a existência de um sistema de recuperação que permite que o usuário encontre mensagens indicando palavras-chave, data, autor, entre outros dados. Esse recurso ajuda consideravelmente alguém que procure uma mensagem específica pois, muitas vezes, o participante não sabe em que

fórum leu determinada mensagem, nem lembra quem a escreveu, mas se recorda de mais ou menos quando ela foi escrita e que palavras ela contém.

3.2. Ambiente Proposto

A análise dos pontos fortes e fracos da ferramenta analisada resultou na idéia de construir um ambiente de comunicação assíncrona que tenha como recurso central uma ferramenta de recuperação para permitir a manipulação de mensagens. Na verdade, essa ferramenta permitiria mais do que a recuperação de mensagens, mas a configuração personalizada de fóruns de discussão.

Esse sistema pode ser visto como um fórum de discussão dinâmico que é configurado a gosto do participante do curso. As mensagens enviadas para o sistema não são vinculadas a um fórum específico, porque o conceito de fórum não faz sentido como uma entidade estática, que é composta de mensagens dispostas por ordem de chegada. Ao invés de visualizar um fórum estático, ao entrar no sistema, o usuário visualizará um configurador para o qual ele dirá que mensagens gostaria de visualizar, fornecendo autor, palavras-chave, data, se ele quer somente as não lidas, e em que ordem devem aparecer.

Nota-se que apenas indicar palavras-chaves é, muitas vezes, insuficiente para que o usuário possa encontrar o conteúdo requerido. Na análise realizada em (Souza & Menezes, 2000a), percebeu-se que é possível definir categorias que agrupam mensagens de conteúdo semelhantes. Com isso, outro ponto central do fórum dinâmico é um categorizador de mensagens, que tem como base categorias criadas pelo administrador do sistema. Essas categorias passam a fazer parte de um catálogo que pode ser consultado pelos administradores dos cursos, para que importem as categorias já construídas.

No fórum dinâmico, não é preciso se preocupar com o tema da mensagem para enviá-la para o fórum específico, pois os temas são dados pelas categorias e os fóruns são construídos dinamicamente, a partir da especificação de, por exemplo, que categorias de mensagens o usuário gostaria de visualizar.

O uso das categorias também fornece mais um recurso na identificação das mensagens para que o usuário não dependa apenas do assunto indicado pelo autor da mensagem para encontrar o conteúdo de que precisa.

É importante destacar que a criação dessas categorias não é uma tarefa muito simples, assim, o ambiente traz alguns recursos para auxiliar o administrador do curso nessa tarefa. Em primeiro lugar, o administrador pode visualizar as palavras e expressões mais utilizadas no curso, para que, a partir desses termos, ele possa criar novas categorias. O sistema verifica também quais as categorias do catálogo utilizam os termos mais freqüentes do curso, para sugerir ao administrador que importe tais categorias.

De tempos em tempos, o sistema reúne as mensagens que não foram classificadas e as apresenta ao administrador, agrupadas por similaridade de conteúdo. Isso também lhe auxilia a criar novas categorias ou refinar as categorias existentes, para que essas mensagens passem a ser classificadas. Para agrupar as mensagens por similaridade, o sistema adapta um algoritmo comumente utilizado na recuperação de documentos em sistemas de RI que utilizam o modelo vetorial, descrito em (Baeza-Yates & Ribeiro-Neto, 1999).

O sistema proposto também fornece um recurso para auxiliar o usuário a encontrar a resposta de uma pergunta. Antes de inserir uma nova pergunta, o sistema verifica se aquela pergunta já foi feita anteriormente. Em caso afirmativo, a resposta da pergunta é retornada para o usuário e, caso ele se satisfaça, a nova mensagem não é

inserida no sistema, evitando o acúmulo de mensagens semelhantes. Para encontrar a resposta mais adequada à pergunta feita pelo usuário, o sistema procura aproximar essa pergunta a outras perguntas já feitas anteriormente, executando o algoritmo utilizado para combinar documentos à *query* do usuário em sistemas que utilizam o modelo vetorial, descrito em (Baeza-Yates & Ribeiro-Neto, 1999).

Experiências que utilizam fóruns de discussão em contextos educacionais (Funaro, 1999) chamam a atenção para o fato de que o uso desse recurso amplia a capacidade de argumentação dos aprendizes, ou seja, eles passam a discutir melhor, analisando o ponto de vista alheio e fornecendo suas próprias idéias sobre um certo tema. Para isso, entretanto, é preciso que os aprendizes leiam inúmeras mensagens e extraíam delas os pontos de vista já levantados na discussão. Essa tarefa pode ser automatizada e é uma das funcionalidades do sistema proposto.

Um outro fator essencial em ambientes que objetivam promover interação mútua, é promover o encontro de usuários afins, tanto com respeito a interesses educacionais como pessoais. Por isso, o sistema proposto inclui também suporte à socialização dos participantes do curso. O sistema mantém um perfil do participante, com suas preferências, suas áreas de interesse, seu grupo de amigos e, além disso, mantém dados estatísticos sobre o grau de interação do participante. Com isso, o sistema sugere ao participante que contacte outros participantes do curso, que tenham perfis semelhantes. Procura-se motivar, especialmente, aqueles que interagem menos a conversarem mais com os demais participantes.

O participante pode acessar, a qualquer momento, o arquivo de FAQ, que é construído dinamicamente, contendo as perguntas mais freqüentes do curso. Uma diferença interessante entre os arquivos de FAQ tradicionais e o gerado pelo sistema é que este reúne as perguntas que realmente foram feitas mais freqüentemente pelos participantes do curso, enquanto aquele reúne as perguntas consideradas mais importantes por quem escreveu o arquivo de FAQ.

Na interação virtual, especialmente em ambientes de aprendizagem, é muito comum que os usuários enviem material uns para os outros. Em grande parte dos sistemas, inclusive na ferramenta analisada, isso se dá pelo envio de arquivos anexados a mensagens, no correio eletrônico e no fórum de discussão. Isso, além de aumentar bastante o tamanho das mensagens, dificulta o acesso posterior ao material. É impressionante dar suporte à organização do material trocado entre os aprendizes, por isso, o sistema permite que o participante publique seu material (textos, fotos ou outras mídias) em um espaço que seja acessado apenas por outros participantes autorizados por ele.

As funcionalidades do ambiente proposto podem ser vistas na figura 1, que apresenta os casos de uso do sistema. Identificou-se dois atores para o sistema: o administrador e o participante do curso. O primeiro é o responsável pela configuração do curso. Quanto ao segundo, preferiu-se utilizar uma definição geral de participante do curso, ao invés de especificar papéis como professor, aluno, orientador, etc., pois deixou-se a cargo do administrador que definisse os papéis relevantes de acordo com as características do seu curso.

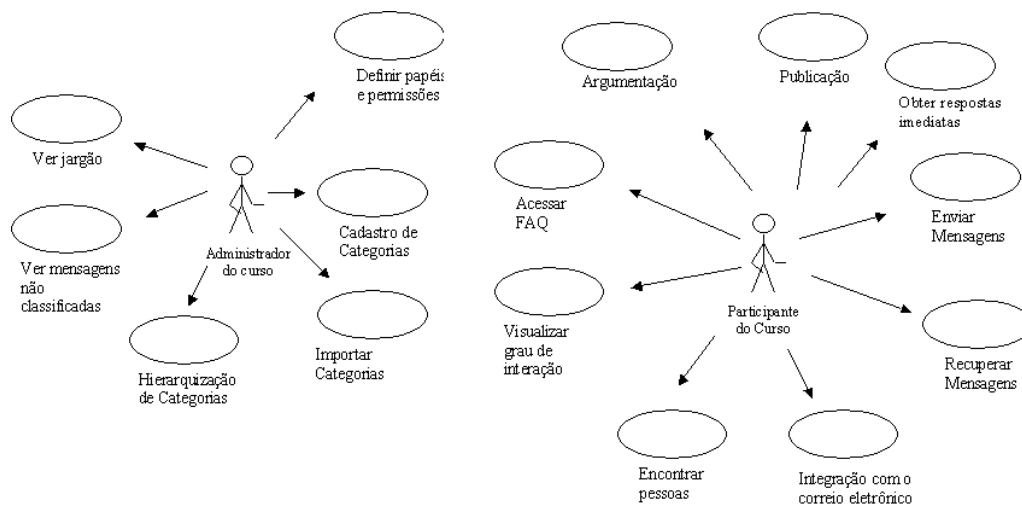


Figura 1 – Casos de uso do sistema proposto

4. Desenvolvimento de um Protótipo

Após a identificação dos casos de uso, passou-se ao desenvolvimento de um protótipo do sistema proposto. O protótipo foi desenvolvido como um sistema multiagentes, em que as funcionalidades são realizadas a partir da cooperação dos agentes construídos.

Para a modelagem, foi utilizada uma metodologia especialmente criada para o desenvolvimento de sistemas multiagentes, denominada GAIA (Wooldridge et. al., 2000). Depois de modelados os agentes, o desenvolvimento prosseguiu com a utilização de um projeto orientado a objetos. Finalmente, o protótipo foi implementado em java, utilizando-se a biblioteca JATLite (JATLite, 2001). Essa biblioteca provê a infraestrutura básica para que os agentes se comuniquem, utilizando a linguagem KQML.

A figura 2 exibe a arquitetura do sistema, mostrando como os agentes se comunicam entre si e como eles se comunicam com o usuário. O usuário interage com o sistema através de um browser, assim, a interface foi implementada por meio de formulários HTML. Entre os formulários e os agentes, existem *servlets*, que intermediam a troca de informações entre o usuário e os agentes. Como essas tarefas exigem uma resposta simultânea, a comunicação é feita de forma síncrona, através da chamada dos métodos dos agentes. Em outras situações, a comunicação entre agentes ocorre de forma assíncrona, por meio da linguagem KQML. Alguns agentes percebem mudanças no ambiente e, então, realizam tarefas, enviando os resultados por email ao usuário. A figura mostra que os agentes estão distribuídos na Internet, alguns deles em servidores HTTP. Todos os agentes se comunicam com o banco de dados, que também está localizado em uma máquina remota na Rede.

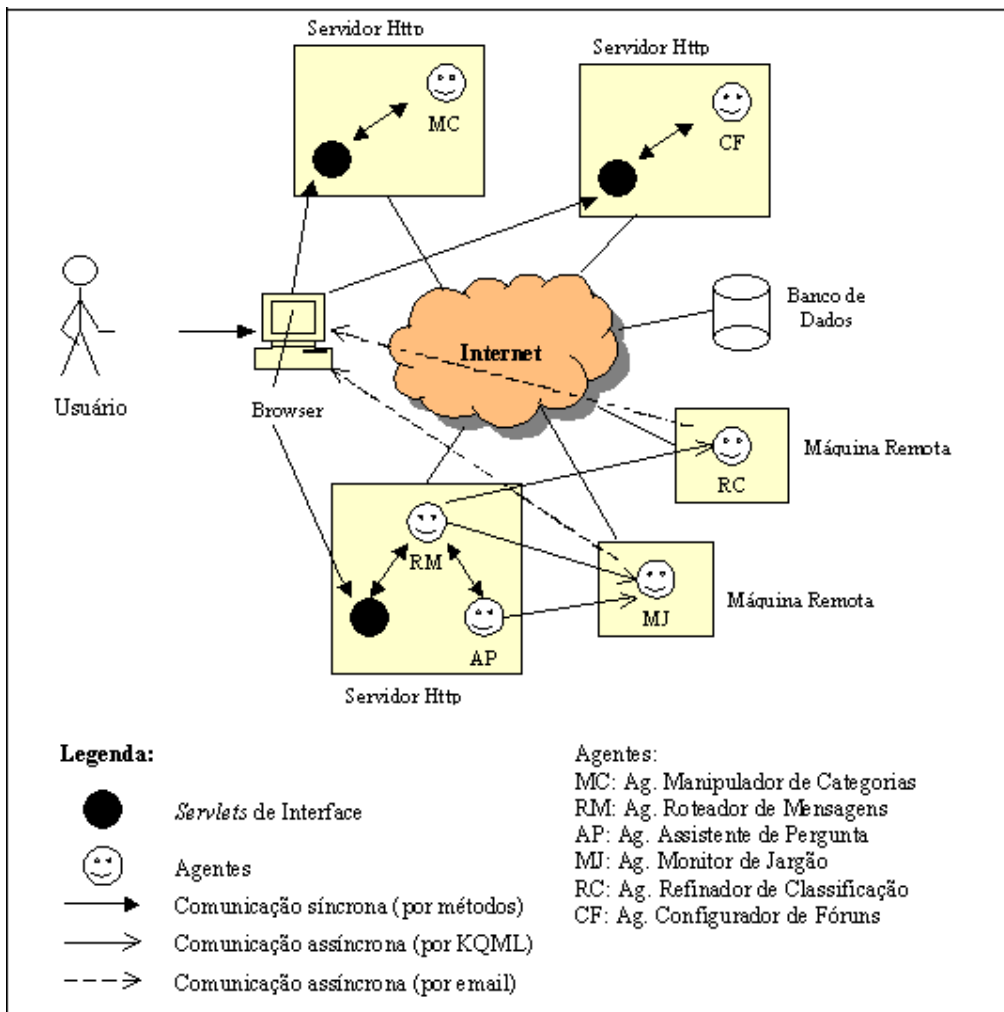


Figura 2 – Arquitetura do sistema

4.1. Recuperação de Mensagens pelo Configurador de Fóruns Dinâmicos

Nos sistemas mais comuns de RI, dos mais antigos aos mais atuais, pode-se distinguir três fases características do processo de recuperação de informações (Salton & McGill, 1983).

- *Análise e Indexação*: cada novo documento é analisado antes de ser inserido na coleção e um conjunto de termos ou identificadores capazes de representar o conteúdo do documento é definido. Esse conjunto de termos é conhecido como *index* do documento.
- *Querying*: consiste na etapa em que o usuário expressa sua necessidade de informação através de uma requisição ao sistema. Para isso, cada sistema disponibiliza para o usuário uma linguagem de *querying*.
- *Recuperação e Apresentação*: a recuperação diz respeito à ação de combinar os termos da *query* do usuário ao *index* previamente estabelecido no sistema, para gerar um conjunto-resultado de documentos que será apresentado ao usuário.

As mensagens enviadas ao sistema proposto são documentos curtos, assim, decidiu-se considerar todas as palavras do conteúdo da mensagem como *index*. Na prática, essa solução era muito cara e consumia muito recurso dos computadores mais antigos, entretanto, hoje, é uma abordagem comum das máquinas de busca da Web (Baeza-Yates, Ribeiro-Neto, 1999).

Como as mensagens estão armazenadas em um banco de dados relacional, a linguagem de *query* adotada foi a própria linguagem SQL. Não é necessário que o usuário conheça essa linguagem, ele expressa sua necessidade de informação através do preenchimento de um formulário HTML. O configurador, por sua vez, traduz as informações inseridas nesse formulário para um comando SQL e faz a consulta no banco de dados, retornando ao usuário as mensagens que satisfazem a sua necessidade.

5. Considerações Finais

Não há dúvidas de que com as ferramentas de comunicação existentes, é possível criar cursos mediados pelo computador que sejam altamente interativos. Para isso, entretanto, é preciso que os participantes do curso, entre eles, professores, alunos, monitores, etc., despendam enorme tempo e esforço na realização de operações para motivar e avaliar a interação, manipular as mensagens do curso, encontrar informações relevantes, organizar o material enviado como arquivos *atachados* às mensagens, entre outras funções.

O ambiente proposto neste trabalho visa a dar suporte a essas atividades, para que essas responsabilidades deixem de ser dos usuários e passem a ser do sistema. Já existem técnicas de Recuperação de Informações que podem ajudar na realização de muitas dessas tarefas. Cabem aos desenvolvedores de software para a educação se apropriarem dessas técnicas para desenvolverem ferramentas mais adequadas à aprendizagem colaborativa.

Este trabalho enfatizou o uso de ferramentas de comunicação assíncrona na aprendizagem, mas nada impede que os mesmos recursos aqui descritos sejam aplicados também a ferramentas de comunicação síncrona, como programas de chat, por exemplo.

Como trabalho futuro, é preciso submeter o protótipo do sistema proposto a uma experiência prática para que se possa analisar os ganhos com o uso do ambiente, bem como avaliar como o protótipo pode ser melhorado.

Referências Bibliográficas

1. BAEZA-YATES, Ricardo; RIBEIRO-NETO, Berthier. **Modern Information Retrieval**, Addison-Wesley, 1999
2. BLATTMANN, Ursula, FACHIN, Gleisy R. B., RADOS, Gregório J. Recuperar a Informação Eletrônica pela Internet. [online] Disponível: <http://www.ced.ufsc.br/~ursula/papers/buscanet.html> [capturado: 15 mai. 01]
3. CASTRO JR, Alberto Nogueira, SILVA, Ana Zuleide Barroso. Um Assistente Inteligente para a Exploração de Recursos Instrucionais. In: WORKSHOP SOBRE AMBIENTES DE APRENDIZAGEM BASEADOS EM AGENTES, 1, Maceió, 2000.
4. FUNARO, G. M. Pedagogical Roles and Implementation Guidelines for Online Communication Tools. **ALN Magazine**, v. 3, n. dez. 1999.
5. GRAVES, Liana N. Creating a Community Context for Cooperative Learning. **Handbook of Cooperative Learning Methods**. Westport: Praeger, 1994. p. 283-299.
6. JATLite [online]. Disponível em: <http://java.stanford.edu>. [capturado em 03 jun. 2001]
7. LOPES, Claudia Mara Amigo, MENEZES, Crediné Silva. Amigo OnLine: Um Ambiente Cooperativo para Recomendações Bibliográficas. In: SIMPÓSIO

- BRASILEIRO DE INFORMÁTICA NA EDUCAÇÃO, 11, **Anais...** Maceió: UFAL, 2000.
8. MOREIRA, Walter. Sistemas de Armazenamento e Recuperação ou Sistemas de Busca: A Recuperação da Informação em Evolução. Anais do I Ciberética - Simpósio Internacional de Propriedade Intelectual, Informação e Ética. Florianópolis, SC, Brasil - 19 a 21 de novembro de 1998. [online] Disponível em: <http://www.ciberetica.iaccess.com.br/anais/doc/sistemasdearmazenamento.doc> [capturado: 15 mai. 01]
 9. SALTON, Gerard, MCGILL, Michael J. **Introduction to Modern Information Retrieval**, McGraw-Hill Book Company, 1983
 10. SOUTO, Maria Aparecida M. S., OLIVEIRA, José P. M., VICCARI, Rosa Maria. Construindo um Ambiente Educacional na Web. In: SIMPÓSIO BRASILEIRO DE INFORMÁTICA NA EDUCAÇÃO, 10., 1999, Curitiba. **Anais...**, Curitiba: UFPR, 1999. p. 295-302.
 11. SOUZA, Renata S., MENEZES, Crediné S. Social Interactions in Cooperative Learning Environments. IN: IEEE ICECE2000 - INTERNATIONAL CONFERENCE ON ENGINEERING AND COMPUTER EDUCATION, **Anais...** São Paulo, 2000.
 12. SOUZA, Renata S., MENEZES, Crediné S. Um Sistema Inteligente para apoio à Interação em Ambientes Cooperativos de Aprendizagem. In: SIMPÓSIO BRASILEIRO DE INFORMÁTICA NA EDUCAÇÃO, 11, **Anais...** Maceió: UFAL, 2000.
 13. WOOLDRIDGE, Michael, JENNINGS, Nicholas R., KINNY, David. The Gaia Methodology for Agent-Oriented Analysis and Design. **Journal of Autonomous Agents and Multi-Agent Systems**, 3, 285-312, 2000.