

Agentes de Software para auxiliar ao Professor na busca de Conteúdos Educacionais no Padrão IMS

Jorge Luis Goñi
jgoni@les.inf.puc-rio.br

Ruy Luiz Milidiú
milidiu@inf.puc-rio.br

PUC-RioInf.MCC13/02 Junho, 2002

Abstract: With the growth of the amount of data available on the Internet, the use of Software Agent for information search and treatment is an option. In Web-Based Education the patterns defined by IMS for interoperability of education contents are a reality. Starting from the experiences and uses demonstrated by the literature, types of agent and their application in the search and treatment of data in the Web are presented. This work also presents a multi-agent system that favor the interoperability of contents in distance education, helping the teacher in the search in IMS-compliant servers.

Keywords: Software Agent, Web-Based Education, Interoperability of Educational Contents, Search engine

Resumo: Com o crescimento da quantidade de dados disponíveis na Internet, a utilização de Agentes de Software para busca e tratamento de informações é uma opção. Na Educação Baseada na Web, existem, por exemplo, os padrões da IMS para a interoperabilidade de conteúdos educacionais. A partir das experiências e utilizações demonstradas pela literatura, são apresentados alguns tipos de agentes e sua aplicação na busca e tratamento de dados na Web. Este trabalho apresenta a utilização de um sistema multi-agentes que apoia a interoperabilidade de conteúdos em educação à distância, fornecendo auxílio ao professor na busca em servidores compatíveis com o padrão IMS.

Palavras Chave: Agente de Software, Educação a Distância Baseada na Web, Interoperabilidade de Conteúdos Educacionais, Motor de Busca

1 Introdução

A extração de conteúdo da Web tem sido objeto de muito interesse por parte da comunidade científica e das empresas de software, tanto no Brasil quanto no contexto internacional. A extração de conteúdo, a partir de páginas na Web, objetiva identificar informações relevantes para um determinado domínio de assunto, acessível a agentes de software e a processos automatizados.

Para apoiar ao professor na busca de conteúdos educacionais utilizou-se a tecnologia de agentes de software. O uso desta tecnologia [1], se apresenta como uma estratégia promissora para ser aplicada aos desafios dos ambientes educacionais modernos que estão cada vez mais influenciados por tecnologias como Internet, Inteligência Artificial e Ferramentas de Busca na Web, consideradas no ContentNet: Um Framework para Interoperabilidade de Conteúdos Educacionais Utilizando o Padrão IMS [2].

Esta monografia apresenta a utilização de sistemas multi-agentes que apoiam os aspectos de interoperabilidade de produtos e conteúdos em educação à distância. No caso do auxílio ao Professor na busca de conteúdos educacionais, propomos como uma solução, a utilização de agentes para possibilitar a reutilização dos conteúdos espalhados em servidores compatíveis com o padrão IMS. A partir das experiências e possíveis utilizações demonstradas pela literatura são apresentadas classificações de agentes disponíveis na literatura a partir das quais são exemplificados alguns tipos de agentes e sua aplicação direta na Web.

Na Seção 2, é resumida a Tecnologia de Agentes. Na Seção 3, destaca-se as Ferramentas de Tecnologia de Busca na Web. Na Seção 4, a descrição dos Conteúdos digitais para educação em Plataforma IMS. Na Seção 5, são analisados os Agentes para pesquisa de conteúdos em Ambiente de Educação, para finalmente propor uma especificação do Sistema.

2 Tecnologia de Agentes

A tecnologia de agentes parece ser uma estratégia promissora para ser aplicada aos desafios atuais dos ambientes educacionais modernos que estão cada vez mais influenciados por tecnologia como Internet, Inteligência Artificial entre outras. Em especial a Inteligência Artificial está sempre provendo mecanismos para serem aplicados às áreas pedagógicas, como por exemplo, os sistemas especialistas de outrora e que ainda têm suas aplicações nos dias atuais. Com o surgimento da Inteligência Artificial Distribuída [3] vivenciou-se e vivencia-se um momento de fortalecimento dos conceitos de agentes de software e também percebe-se um redirecionamento destes conceitos para a Engenharia de Software [4].

Segundo [8], as aplicações que utilizam agentes ainda são limitadas. E uma vez que os conceitos se tornem mais aceitos, a abordagem de agentes vai se tornar mais difundida nas aplicações de tecnologia da informação. Um exemplo, dentro das categorias de agentes, as

possíveis utilizações de *agentes de interface* dentro de um ambiente de educação baseada na Web.

2.1 Agentes de Software

O conceito de agentes de software ainda não está consolidado. Em diferentes grupos de pesquisa, para se definir um agente, diferentes características podem se fazer necessárias. Isso leva a uma grande variedade de definições e características, o que foge do escopo deste trabalho. Uma visão geral sobre este assunto pode ser obtida em [5] e [6].

Assim como existem várias definições para agentes, como por exemplo, “Um agente é uma entidade de software que funciona de forma contínua e autônoma em um ambiente em particular, geralmente habitado por outros agentes, e que seja capaz de intervir no seu ambiente de forma flexível e inteligente, sem requerer intervenção ou orientação humana constante. De um modo ideal, que funcione continuamente por longos períodos de tempo.” [7], também existem várias taxionomias. Segundo [1], existem sete categorias de agentes: colaborativos, de interface, móveis, de informação, reativos, híbridos e “inteligentes”. Em [8], a apresentação de uma série de aplicações atuais utilizando agentes pode ser encarada como uma classificação dos tipos de agentes encontrados. Existem várias outras classificações, por exemplo em [5], novas categorias são apresentadas ou apenas reestruturadas baseando-se em [1]. A seguir, apresenta-se uma revisão das categorias encontradas em [1] e [8] através da identificação de (possíveis) implementações de agentes na área educacional, especificamente para instrução baseada na web.

Segundo [1], um agente pode ser classificado em um ou mais dos tipos descritos a seguir.

Agentes inteligentes: este tipo de agente não é definido pelo autor, pois segundo ele este tipo é mais uma aspiração dos pesquisadores do que realidade;

Agentes de informação: são responsáveis pelo gerenciamento, manipulação e ordenação de informação de diferentes fontes. São caracterizados por aquilo que “fazem” e não por aquilo que “são”;

Agentes móveis: são processos de software capazes de se locomover em uma rede de computadores. “Contudo, mobilidade não é uma condição necessária nem suficiente para caracterizar um agente.” [1];

Agentes reativos: são os agentes capazes de reagir a estímulos do ambiente em que se encontram, mas não possuem modelos internos do mesmo. Um aspecto importante é que a partir de interações (respostas) simples, surgem padrões de comportamento complexos, quando analisamos o sistema de agentes reativos;

Agentes híbridos: são aqueles constituídos através da combinação de duas ou mais “filosofias” de agente;

Agentes de interface: colaboram com o usuário em um ambiente de trabalho, desempenhando o papel de assistentes pessoais;

Agentes colaborativos: colaboram com outros agentes para realizar determinadas tarefas.

Os agentes podem apresentar características tais como aprendizagem, cooperação e mobilidade, então podemos definir um agente como uma entidade computacional que:

- age no lugar de outra entidade de uma forma autônoma;
- realiza suas ações com um nível de pró-atividade e/ou reatividade;
- possui um nível de aprendizagem, cooperação e mobilidade.

2.2 Sistemas Multi-Agentes

A partir da década de 1970, com o surgimento dos Sistemas Tutores Inteligentes, os pesquisadores em Informática na Educação observaram a necessidade de usar técnicas de Inteligência Artificial para tornar os sistemas de ensino mais flexível e adaptados aos seus usuários. Atualmente, o enfoque colaborativo dado ao ensino à distância desencadeou um grande número de problemas e atividades em que a tecnologia de agentes pode ser muito bem empregada, tanto auxiliando e monitorando os alunos como também provendo informações ao professor.

Quando um sistema é formado por mais de um agente, ele é conhecido como Sistema Multi-Agentes (MAS). Segundo [9], “é uma rede fracamente acoplada de solucionadores de problemas que trabalham em conjunto para resolver problemas que vão além da sua capacidade individual. Estes solucionadores de problemas são essencialmente autônomos, distribuídos e, muitas vezes heterogêneos em sua natureza.”

Os agentes são entidades autônomas que têm conhecimento da sua própria existência e da existência de outros agentes e, portanto, colaboram uns com os outros para atingirem um objetivo comum dentro de um ambiente.

Em um MAS, os agentes devem possuir algumas capacidades específicas para interagirem num mesmo ambiente. Portanto, os agentes devem ter conhecimento da sua existência e da existência dos outros agentes. Além disso devem ser capazes de se comunicar possuindo, para tanto, uma linguagem específica. Cada agente deverá possuir conhecimentos e habilidades para executar uma determinada tarefa e devem cooperar para atingir um objetivo global.

Para o MAS resolver problemas coerentemente, os agentes devem:

- comunicar entre si;
- coordenar suas atividades e
- negociar em caso de conflitos.

Os conflitos entre os agentes podem variar de simples retenção de recursos limitados à computações complexas onde os agentes discordam por causa de discrepância entre seus domínios de especialidade.

A coordenação é fundamental no MAS pois sem ela, qualquer benefício de interação desapareceria e o grupo de agentes rapidamente degenerar-se-ia numa coleção de indivíduos com um comportamento caótico.

A interação entre um grupo de agentes pressupõe:

- alocação de recursos e tarefas;
- negociação;
- detecção e resolução de conflitos.

Para armazenar o conhecimento no caso dos agentes inteligentes é necessário definir uma forma de representação de conhecimento, que é definida como a combinação de estruturas de dados e procedimentos interpretativos, que se usados de maneira adequada levam um agente a ter um comportamento instruído (possui conhecimento).

As técnicas de representação de conhecimento [10] são diversas. A seguir, são enumeradas e descritas, algumas delas.

Direta ou Analógica: representação de propriedades e relações entre as partes (mapas, geometria, diagramas, etc.).

Lógica: usada para representar relações e estruturas dedutivas. Aqui, uma determinada declaração pode ser verdadeira ou falsa (lógica de predicados).

Regras de produção: sistema de representação baseado em regras do tipo se <condição> então <ação>. É dividido em três partes: base de conhecimento na forma de regras, contexto (foco de atenção das regras a serem utilizadas) e interpretador (decide o que fazer).

Procedimentos: procedimentos de análise sintática (a gramática não é explícita).

Rede semântica: estrutura envolvendo nós (representando objetos, conceitos ou ações) e arcos (ligando os nós e representando relações entre eles).

Frames: pacote de dados transmitido incluindo informações de controle e rota.

Scripts: facilita a criação de páginas dinâmicas na Web.

Embora essa seja ainda uma área bem atual, já se encontram alguns exemplos de aplicações usando agentes. O OMG [8], por exemplo, apresenta uma série de aplicações utilizando agentes que pode ser encarada como uma classificação dos tipos de agentes encontrados atualmente. A seguir, apresentamos alguns exemplos de tipos de aplicações possíveis.

Aplicações empresariais: documentos inteligentes que sabem quando devem ser processados, e gerenciamento dinâmico de pessoal e suas funções.

Aplicações Business-to-business: criação de mercado para bens e serviços, e gerenciamento de equipes.

Controle de processos: gerenciamento de fábricas (refinarias), e prédios inteligentes (aquecimento/resfriamento inteligente, segurança inteligente).

Computação móvel: provêem aos desenvolvedores de aplicações distribuídas um conjunto de abstrações que mascarem a complexidade e heterogeneidade normalmente encontradas em infra-estruturas distribuídas.

Agentes pessoais: filtros de *e-mail*, e gerenciamento de agendas pessoais.

Tarefas de gerenciamento de informação: procura por informação, filtragem, monitoramento de informação entre outros.

3 Ferramentas de Tecnologia de Busca na Web

A web é uma imensa base de dados, uma formidável coleção de textos, livros, artigos, revistas, jornais, e-zines, catálogos de produtos e de serviços, músicas, sons, imagens, fotografias, programas de computador, vídeos, filmes, peças publicitárias. Tudo isso encontra-se por lá. E está ao alcance de todos, desde que se saiba como encontrar o que lhe interessa utilizando os sistemas de busca.

Sem as máquinas de busca seria praticamente impossível encontrar informação na Internet, uma vez que há mais de um bilhão de páginas espalhadas em todo o mundo. O projeto IMS (“*Instructional Management Systems Global Learning Consortium*”), vem desenvolvendo e promovendo especificações para possibilitar a busca e a troca de conteúdos educacionais espelhados em diversos servidores de conteúdo; através da utilização de motores de busca.

Nas ferramentas de Tecnologia de Busca na Web, Sistema de busca é o termo genérico que engloba duas categorias: os motores de busca e os diretórios. O termo "motor de busca" é freqüentemente usado para descrever motores de busca e diretórios. Mas eles não são iguais. A diferença está em como as listas são compiladas.

3.1 Motores de busca

Motores de busca são ferramentas da Web que, ao contrário dos diretórios, utilizam softwares ou programas de pesquisa para efetuarem os processos de busca, indexação e atualização de suas bases de dados, ou seja são baseados no uso exclusivo de programas de computador para a indexação das páginas da web. Os mecanismos de busca têm três componentes principais, segundo [11], as tarefas a serem executadas por estes programas são mostradas esquematicamente na figura 1.

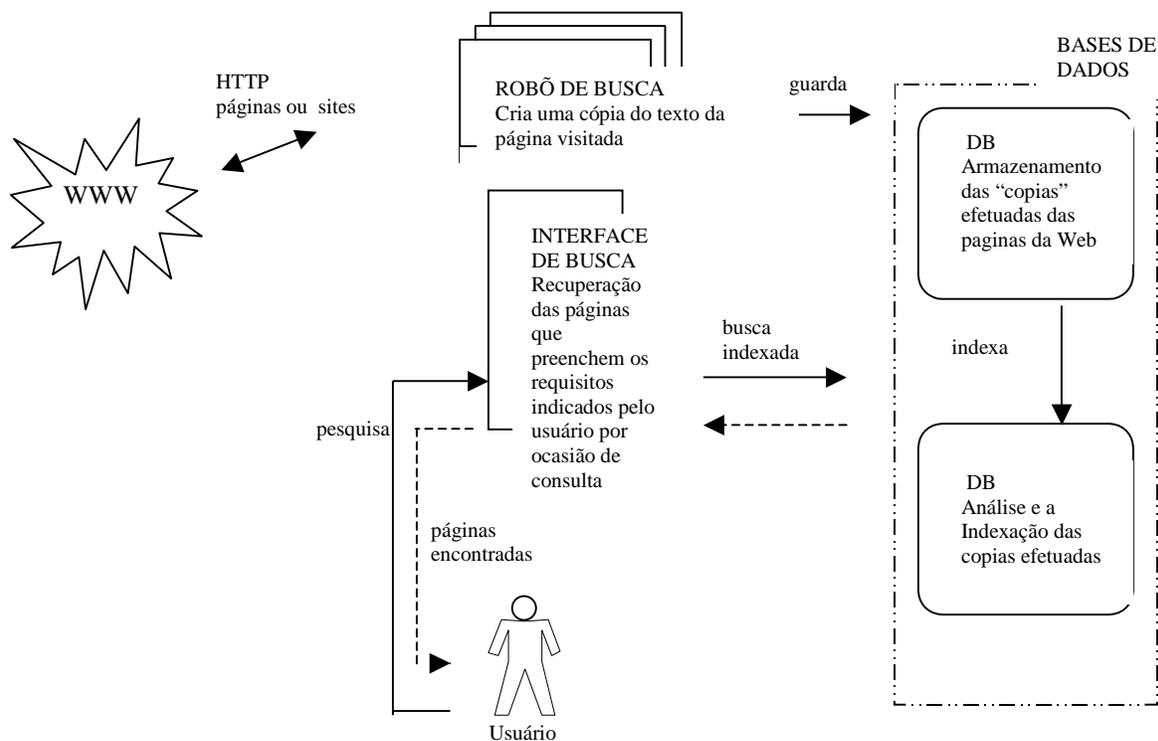


Figura 1. Esquema de um motor de busca genérico para a Web.

Robô de busca

O componente de Software é denominado como robô, *crawler*, *spider*, *wanderer*, *knowbot*, *worm*, *web-bot* ou agentes. Os robôs, varrem a Web, visitando páginas, lendo-as, extraíndo os itens para ele interessantes, movendo-se de um documento a outro por referências de *hiperlinks* embutidos nas páginas e utilizando o protocolo HTTP para recuperar documentos dos servidores. Eles utilizam técnicas heurísticas com o objetivo de encontrar e indexar os sites mais populares primeiro.

Quando os agentes recuperam uma URL, podem aprender a revisitar freqüentemente alguns sites, principalmente aqueles que mudam muito, ou que tem links para muitas páginas. A estratégia dos robôs é seguir as complexas ligações dos hipertextos e “visitar” as páginas ou sites armazenadas na WWW, coletando dados importantes, ou seja apenas as visitam com o objetivo de coletar dados para a indexação destas em suas bases de dados.

A maioria dos robôs possui um período de tempo predeterminado para revisitar os sites por eles indexados visando detectar as mudanças ocorridas naqueles sites. Nos motores de busca, não existe intervenção humana nos processos acima citados. Os critérios de busca, indexação e classificação além de totalmente automáticos são específicos para cada motor.

Os critérios utilizados pelos robôs para efeito de visitação – *crawling* – indexação de novas páginas e itens a serem indexados variam de motor para motor. Alguns robôs utilizam *deep crawl*, um recurso que permite que o robô, acesse muitas páginas de um site, mesmo quando estas não são submetidas a eles. Alguns robôs utilizam visitação em largura,

na qual uma vez que os mesmos se encontram em uma página o robô visita todos os links existentes na mesma.

Bases de dados

O componente base de dados, fica armazenado no servidor do mecanismo de busca e é constituído das cópias ou réplicas efetuadas pelo robô. Os itens coletados pelo robô durante o processo de indexação são armazenados nas bases de dados. Sendo assim o conteúdo destas também é completamente dependente da estratégia de indexação utilizada pelo robô de busca. Exemplos de itens que podem ser encontrados nas bases de dados são endereços ou URL's, títulos, cabeçalhos, resumos, palavras da primeira ou primeiras linhas dentre outras.

O tamanho da base de dados varia de motor para motor e na maioria das vezes, o conteúdo das bases de dados diferem em detrimento da estratégia de indexação individual utilizada por cada mecanismo, levando a uma pequena interseção dos dados contidos nas bases de dados de motores distintos. Tal situação incentiva a utilização de meta-motores, a fim de ampliar a cobertura em uma busca.

O tamanho da base de dados é responsável pela amplitude da pesquisa. Quanto mais links o serviço tiver , mais itens ele recuperará e assim se tornará mais popular.

Existem três maneiras de medir o tamanho das bases de dados destes motores: o número de páginas recuperadas, o número de URL's únicas e o número de URL's.

Software de interface

O último componente é o software de interface que faz a interação do usuário com o motor de busca. É através dele que o usuário fornece suas palavras-chave para serem pesquisadas pelo motor. Após a entrada dos dados, o software analisa minuciosamente as milhões de páginas registradas em suas bases de dados a fim de encontrar aquelas que satisfazem as palavras fornecidas. Em seguida, devolve para os usuários os itens selecionados que satisfazem a busca utilizando para isso um algoritmo de classificação que tenta determinar a ordem de relevância dos documentos recuperados. O resultado deste algoritmo é a apresentação dos itens ordenados por ordem de relevância. Assim, o primeiro item listado pelo motor de busca é o que conseguiu maior peso na estratégia de classificação utilizada pelo mecanismo. A experiência porém detecta que nem sempre o julgamento feito pelo algoritmo de classificação coincide com o do usuário.

Dentre os principais motores de busca nacionais, destacamos:

MSN Brasil - <http://www.msn.com.br/>

Uma particularidade deste site é que ele permite a busca por tipo de arquivo (vídeo, Active x, Shockwave, JavaScript, áudio, áudio mp3, VBScript, Acrobat, miniaplicativos Java)

Altavista Brasil - <http://altavistabrasil.com.br/>

- <http://br.altavista.com/>

Portal de buscas americano, oitavo colocado no ranking global da Media Metrix, que entrou no mercado brasileiro em dezembro de 2001. Em sua

versão original, incluía 8 milhões de páginas da web brasileira, oferecia opções de pesquisa avançada, busca multimídia e tradução online. Na seção "Minha Pesquisa", os usuários podem armazenar até 25 buscas. Também permite personalizar os resultados das pesquisas.

Lycos - <http://www.lycos.com.br/>

Indica os 50 sites mais procurados (ver www.lycos.com.br/50)

Dentre os principais motores de busca estrangeiros, destacamos:

Google – <http://www.google.com/>

Atualmente, o maior motor de busca no mundo. Indexa 1,5 bilhões de páginas.

Altavista – <http://www.altavista.com/>

Um dos maiores motores de busca no mundo. Indexa 550 bilhões de páginas.

Fast Search – <http://www.alltheweb.com/>

Lançado em maio/1999. Indexa 575 milhões de páginas. Norueguês.

Northern Light – <http://www.northernlight.com/>

Indexa 350 milhões de páginas. Motor de busca que se destaca pela categorização dos resultados por conceitos relacionados. Oferece também pesquisa em bases de dados que cobrem mais de 7,000 periódicos, com artigos desde 1995. Paga-se para ver o conteúdo dos artigos.

AOL Search – <http://search.aol.com/>

Permite visualizar resultados por popularidade dos links. Clicar em "Most popular sites".

WISEnut - <http://www.wisenut.com/>

Novo motor lançado em out/2001. Provável competidor para Google.

3.2 Diretórios

O Diretórios são os sistemas de busca nos quais a indexação das páginas da web é realizada por humanos.

Um diretório como o Yahoo [13], depende dos humanos para suas listas. Você envia uma pequena descrição para o diretório da sua página. Uma busca procura somente por referências nas descrições enviadas.

Mudanças nas páginas não afetam como elas são listadas. Métodos que são usados para melhorar a listagem com motores de busca são diferentes dos adotados para uma listagem em diretórios. A única exceção é que um site bom, com bom conteúdo, pode ser visto mais vezes do que um com pouca qualidade.

Dentre os principais motores de busca nacionais, destacamos:

Cade? – <http://www.cade.com.br/>

Diretório de sites mais antigo do Brasil. Pertencente à Starmedia, em dezembro/2000 fundiu-se com o diretório Zeek. Comprado pelo Yahoo em jan/2002.

Busca Brasil – <http://www.brasil.com.br/>
*19 guias segmentados e mais de 190 categorias. Contém notícias, e eventos.
Lançado em 1998.*

Surf – <http://www.surf.com.br/>
Mecanismo que assumiu as informações contidas no antigo Yahi!

Dentre os principais motores de busca internacionais, destacamos:

LookSmart – <http://www.looksmart.com/>
Um dos maiores diretórios do mundo. Afirma, em setembro/2000 ter atingido 2 milhões de links coletados.

Open Directory – <http://dmoz.org/>
Um dos maiores diretórios do mundo. Afirma, em setembro/2000 ter atingido 2 milhões de links coletados

3. 3 Problemas encontrados no sites de busca

A qualidade do serviço oferecido pelos sites de busca depende basicamente do banco de dados disponibilizado e do poder de procura oferecido. Segundo [14] os maiores problemas encontrados são:

1. encontrar somente o assunto desejado dado o grande volume de informação;
2. retorno de fontes inexistentes: em alguns casos a fonte apresentada não existe mais pois pode ter sido tirada da Internet ou movida para outro lugar;
3. dificuldade de realizar pesquisas complexas e dentro de determinado contexto, por exemplo, ao procurar sistemas de agentes não quero que sejam apresentados agentes de viagens ou qualquer outro tipo de agente que não é de software;
4. as buscas são realizadas somente em páginas *html*, não sendo possível definir outros tipos de fontes de dados como banco de dados, diretórios de *ftp*, arquivos locais, etc
5. não existem mecanismos de recuperar as últimas pesquisas realizadas ou aproveitar pesquisas que foram realizadas por outras pessoas sobre o mesmo assunto; e
6. o usuário tem que guardar as suas consultas e se quiser saber se novas fontes de informação foram inseridas na Internet deve realizar a consulta novamente.

4 Conteúdos digitais para educação

O IMS [12] tem como objetivo o desenvolvimento e promoção de especificações abertas para facilitar a distribuição de atividades educacionais como localizar e utilizar conteúdos educacionais: de acompanhamento, de progresso, de informe, de desempenho e das experiências do aprendiz entre sistemas administrativos e gerenciais.

O IMS tem dois objetivos principais, a saber:

- Definir o padrão tecnológico usado para interoperabilidade de aplicações e serviços em aprendizado distribuído, e;

- Dar suporte a incorporações de especificações IMS em produtos e servidores voltados para a Web. IMS espera promover uma especificação adotada por todos, o que possibilitará que ambientes e conteúdos distribuídos, e produzidos por diferentes atores trabalhem juntos, ou melhor, iteroperem.

Desse intento, surgiram vários comitês técnicos, cada qual com uma missão específica de estabelecimento de padrões para a área específica. A seguir são apresentados de forma sucinta, propostas dos grupos do IMS, com seus respectivos temas e sua idéia geral.

IMS Learning Resources Metadata Specifications - cria uma maneira uniforme de descrever recursos educacionais, tal que estes possam ser encontrados mais facilmente usando uma ferramenta de busca baseada em meta-informação;

IMS Enterprise Specification - visa a aplicações e serviços administrativos que precisam trocar informação sobre aprendizes, cursos, performance, etc. Entre plataformas, sistemas operacionais e outros;

IMS Content & Packaging Specification - descreve uma estrutura de dados que será usada para prover interoperabilidade de conteúdos baseados na Web com ferramentas de criação de conteúdos, software de gerenciamento educacional e ambientes educacionais;

IMS Question & Test Specification - descreve uma estrutura de dados que será usada para prover interoperabilidade entre sistemas que utilizam questionários e testes;

IMS Learner Profiles Specification - focaliza-se em maneiras de organizar informações educacionais, tais que sistemas educacionais possam ser mais centrados em suas necessidades específicas.

A padronização *IMS Learning Resources Metadata Specifications* descreve os nomes, definições, organização e as construções dos elementos de metadados IMS e na descrição do próprio conteúdo a ser trocado, não estabelecendo, assim, o protocolo que viabilizará a interoperabilidade dos mesmos. A construção deste documento foi baseada na versão 3.5 do padrão proposto pelo grupo de trabalho P1484.12 *Learning Objects Metadata WG do IEEE Learning Technology Standards Committee (LTSC)*, no qual foram propostas diversas modificações sobre este padrão proposto pelo IMS [15].

Visando à interoperabilidade de conteúdos educacionais, “*IMS Learning Resources Metadata Specifications*” utiliza a descrição de metadados proposta no projeto IMS para possibilitar a busca e a troca de conteúdos educacionais espelhados em diversos servidores de conteúdo. Como mostra a arquitetura na Figura 2.

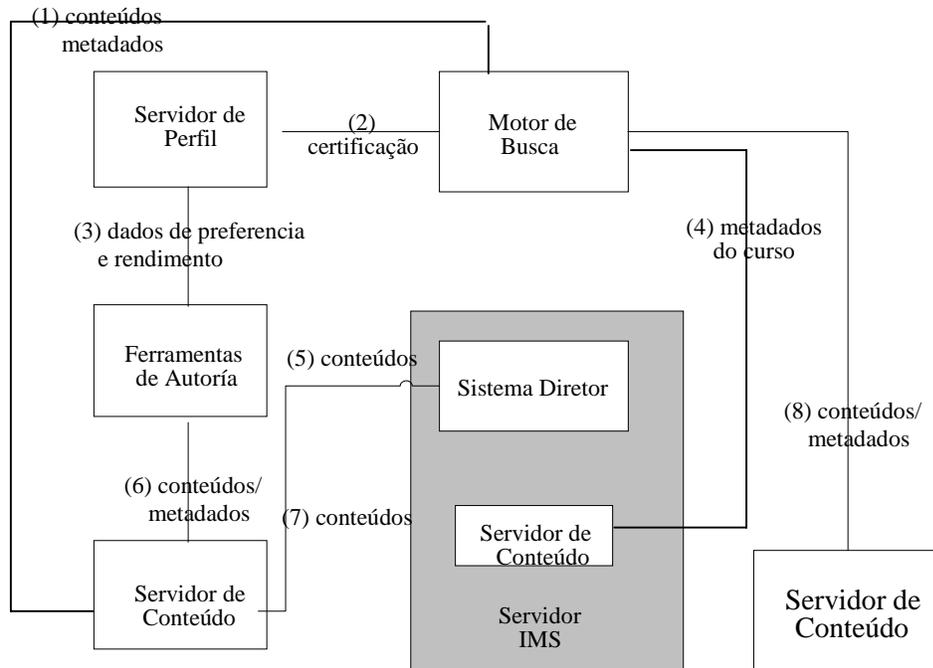


Figura 2. Arquitetura IMS para conteúdos educacionais distribuídos.

Na arquitetura proposta temos os seguintes elementos:

- (1) um motor de busca pode usar a informação dos metadados para perguntar a um servidor de conteúdos sobre materiais de aprendizagem específicos.
- (2) um motor de busca pode perguntar a um servidor de perfis para encontrar aquelas pessoas que desejam certificação de habilidades.
- (3) uma ferramenta de autoria pode perguntar a um servidor de perfis os dados de preferência e rendimento para personalizar sua apresentação.
- (4) um motor de busca pode obter metadado acerca de um curso desde o servidor de conteúdo de um servidor IMS.
- (5) uma ferramenta de autoria pode trocar conteúdo com um sistema gestor IMS.
- (6) uma ferramenta de autoria pode interatuar com um servidor de conteúdo para encontrar conteúdos ou metadados.
- (7) um servidor de conteúdo pode prover conteúdos a um Sistema IMS
- (8) haverá muitos servidores de conteúdos ao alcance de motores de busca.

O conceito de metadado a seguir, é descrito num resumo utilizado no padrão proposto pelo IMS.

Metadados

Os Metadados são organizados em categorias ou campos. Estes campos representam as características dos recursos físicos ou eletrônicos. Cada recurso, por sua vez, é descrito de acordo com a informação contida na combinação destes campos. IMS tem definido uma coleção extensível de metadados para recursos educativos e de treinamento incluindo,

conteúdo, itens de rendimento, resultados de desempenho, informação pessoal, e informações da certificação.

Conteúdo

Pode ser um objetivo educativo ou integrações de objetivos. Podem ser um módulo educativo ou agregações de módulos que tem uma intenção educativo. Pelo outro lado, o conteúdo podem ser algo tão simples como uma palavra ou um quadro.

Sistemas de gestão

O sistema de gestão implementa os serviços que os conteúdos usam.

Perfis

São coleções de dados pessoais e educacionais dos usuários. Um perfil IMS para um usuário pode incluir informação pessoal, de rendimento e preferencias, específica do aluno e/ou específica do professor.

Assim, o modelo IMS oferece um conjunto de ferramentas que facilitam a especificação de um serviço de busca para possibilitar a troca de conteúdos educacionais espalhados em diversos servidores de conteúdo, ou numa federação de servidores, tornando assim possível produzir materiais educativos de alta qualidade e potencialmente reutilizáveis. A plataforma IMS pretende gerar uma infra-estrutura de forma a que toda a infra-estrutura já criada possa ser reutilizada na produção de ambientes de aprendizagem. Esta reutilização tem um valor suplementar frente aos investimentos já realizados, por muitas escolas e universidades, no desenvolvimento de ambientes de aprendizagem.

5 Agentes de busca para Ambiente de Educação a Distância

Agentes que trabalham com busca e tratamento de informação são uma opção para a implementação de serviços apoiados por padrões de interoperabilidade de produtos e conteúdos educacionais [2] como aqueles propostos pelo IMS. São exemplos de agentes nas tarefas de gerenciamento de informação:

- Agentes para procura por informação que saibam como localizar os dados, que saibam como obter acesso ou mesmo negociar os modos de acesso, e que tenham informações sobre o cuidado da fonte;
- Agentes para filtragem de informação que possuam modelos de seus usuários, propiciando uma busca mais efetiva e personalizada;
- Agentes mediadores que facilitem a negociação de diferentes agentes interessados em obter informações ou conteúdos educacionais;
- Agentes capazes de monitorar dados em diferentes fontes de informação sejam elas servidoras de conteúdos educacionais ou não, apoiando de alguma forma a mobilidade dos aprendizes, professores ou projetistas.

Com o crescimento da quantidade de dados disponíveis, a utilização de agentes para busca de informações já vem sendo considerada [16] e a sua aplicação em interoperabilidade de conteúdos educacionais pode sustentar a utilização destes padrões.

5.1 Elementos principais dos motores da busca

A diferença entre o serviço de busca baseado em agentes e um serviço de busca tradicional, está na forma de realizar a busca por informações. No serviço baseado em agentes o usuário pode definir diversos agentes que realizarão o trabalho de buscar de informações em diversos tipos de fontes de dados, como mostra a figura 3. Além disso, o usuário pode definir o seu perfil de pesquisador, através da definição de assuntos de seu interesse em conjunto com sinônimos desses assuntos. Por exemplo, no caso de interesse sobre agentes, o usuário pode definir como sinônimos agente de software, *sistemas de agentes, agent, software agent, etc.*

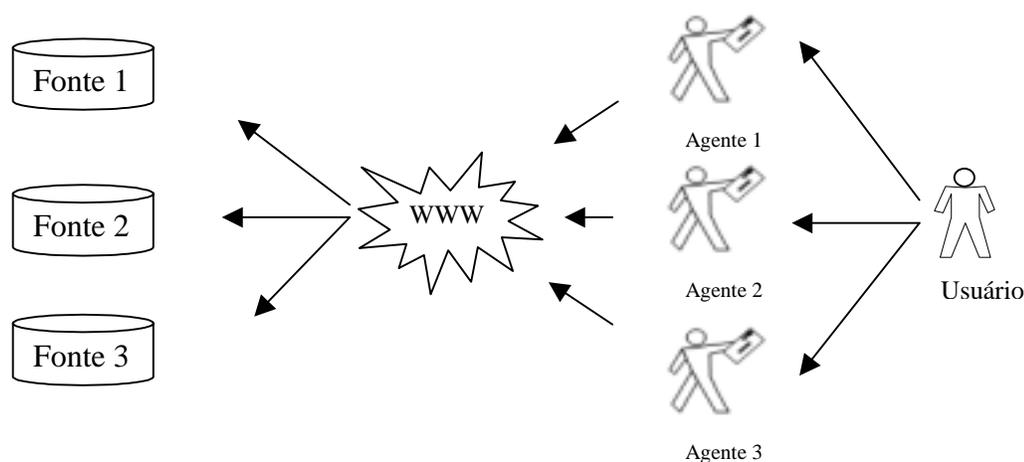


Figura 3. Agentes realizam o trabalho de busca nas diversas fontes de informação.

A tarefa dos motores da busca atenderam os três elementos principais seguintes:

O primeiro é o *spider* (aranha), chamado também de esteira rolante. A aranha visita a página, lê, e segue então os links dentro da página. Isto é o que ocorre quando alguém consulta a um local que é "spidered" ou "rastreado." A aranha retorna ao local em uma base regular, tal como a cada mês ou dois, para procurar por mudanças.

Tudo que a aranha encontra vai para a segunda parte do motor de busca, o index. O index, algumas vezes chamado de catálogo, é como um livro gigante contendo uma cópia de todas as páginas que a aranha encontrou. Se uma página muda, então este livro é atualizado com a nova informação.

Algumas vezes pode decorrer algum tempo até que novas páginas ou mudanças que a aranha encontrou sejam incorporadas ao índice. Assim uma página pode ter sido rastreada mas não indexada. Até que ela seja indexada, incorporada ao index, ela não estará disponível para a procura nos mecanismos de busca.

Os programas de busca são a terceira parte dos motores de busca. Este programa é o que percorre milhões de páginas gravadas no índice para encontrar a referência para a procura e retornar em ordem aquelas que ele acredita serem mais relevantes.

5.2 Fontes de Informação

A Internet permitiu o acesso aos diversos tipos de fontes de informação, apesar da grande parte das informações estarem disponíveis através de páginas HTML, existem muitos banco de dados que armazenam um conjunto imenso de informação e podem ser acessados via Internet. Além disso, diversos tipos de fontes de informação estão disponíveis para acesso. Acreditamos que fazer uso desta fontes pode ajudar bastante na busca de informações, visto que a busca textual é limitada e ineficiente e por outro lado uma busca em um banco de dados apresenta ferramentas que podem facilitar na recuperação de informações.

No serviço proposto, o agente poderá acessar diversas fontes de dados e não somente páginas HTML. As fontes podem ser definidas pelo usuário e o agente poderá fazer uso delas para realizar suas pesquisas, como mostra a figura 4.

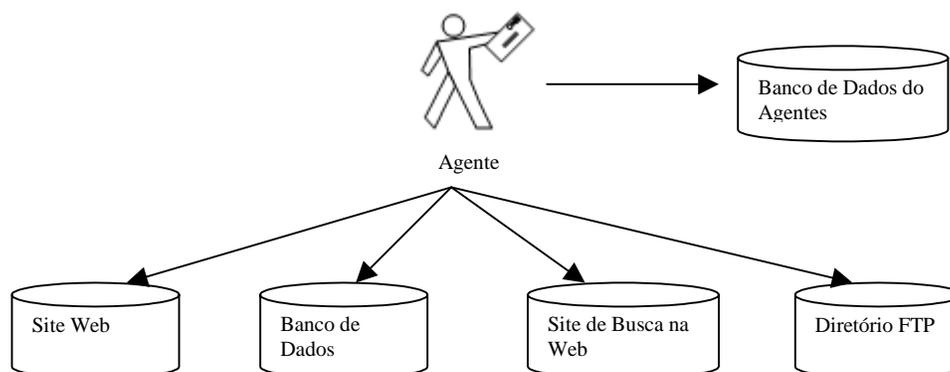


Figura 4. O agente acessa diversas fontes de dados.

5.3 Espaço de Agente

Os agentes criados realizarão as atividades delegadas pelo usuário sem a necessidade de estímulo externo, isto significa que após a saída do usuário do serviço, os seus agentes continuarão ativos até que o usuário os interrompa. O agentes criados por diferentes usuários podem trocar informações entre si. Essas informações podem ser as pesquisas realizadas, as fontes pesquisadas e detalhes das buscas realizadas, como tempo de procura, itens recuperados, etc. Para permitir a atuação dos agentes e a interação entre eles é necessária a construção de um local onde eles possam se comunicar. Esse local é denominado de espaço de agentes e segue o modelo de “*blackboard*” para interação entre agentes, como mostra a figura 5. Os agentes criados serão persistidos em um banco de dados relacional e poderão ser ativados ou desativados pelo espaço de agentes de acordo com a sua necessidade de execução.

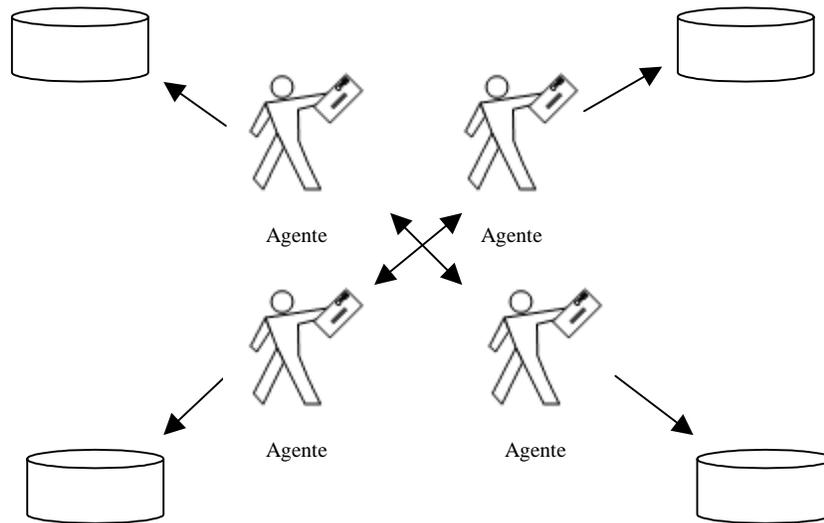


Figura 5. Os agentes de diferentes usuários trocam informações entre si.

5.4 O Agente de Pesquisa

A função do agente é procurar informações em nome de seu usuário e informá-lo de suas descobertas. O usuário pode definir:

- os assuntos que o agente deve procurar e as fontes onde ele deve buscá-las,
- parâmetros para sensibilidade do agente (tempo de execução) e a forma de informar ao usuário sobre as informações encontradas,
- consultas através de palavras-chave, utilizando operadores do tipo e, ou, não, perto, etc.

As consultas do usuário são realizadas sobre o banco de dados e as informações são recuperadas para usuário informando a fonte de informação que contém a palavra-chave desejada.

5.5 Agentes de Pesquisa em Ambiente de Educação a Distância

Na Web é fácil colocar informações porém difícil buscá-la, acessar e obter essas informações devido ao grande número de recursos disponíveis.

A utilização da Web no ensino a distância tem levado à mudança de paradigma da forma tradicional de ensinar. Hoje estudantes e professores, podem ter acesso através da Web as mais variadas formas de informações distribuídas como: bibliotecas digitais, artigos de congressos, livros digitais, entre outros. Além da troca de informações entre as pessoas geograficamente distribuídas, a Web propicia intensa iteração na reutilização de conteúdos educacionais contidas em servidores.

5.5.1 Especificação do Sistema

O sistema proposto descreve o uso de agentes de Software para auxiliar a tarefa do professor, na busca de conteúdos educacionais disponíveis em servidores do ambiente de Educação a Distância baseado na Web, facilitando seu acesso, descrição, localização e uso que permita sua reutilização.

O sistema é formado por quatro tipos de agentes, como mostra a figura 6, um agente do professor, que propicia a interface do professor na busca de conteúdos educacionais e três agentes coletores, programas que varrem os servidores para buscar de informação.

Agente do professor

Este agente, chamado *agente monitor*, é responsável por exibir as análises feitas pelos outros agentes.

Os agentes coletores

Estes agentes são responsáveis por coletar dados a partir dos metadados armazenados nos conteúdos dos arquivos do curso, onde:

- os metadados se referem a dados sobre dados ou informação sobre os recursos didáticos disponíveis, por exemplo: Fichas de bibliotecas contêm entradas de metadados para livros e outros recursos encontrados em uma biblioteca [17]. O conjunto dos metadados associados às características de um livro seriam: *Título, descrição, autor, formato, nível educacional, versão, etc.*
- um conteúdo é composto de um ou mais arquivos, descrito por elemento como por exemplo, um, texto, uma figura, um áudio, um vídeo. Um elemento é tipicamente, um arquivo com o formato *GIF, JPG, HTML, TXT, PDF, e etc.*

No sistema o professor poderá fazer uma busca através de um agente monitor, dos arquivos que são colocados no servidor de ambiente que deveram ser registrados com suas palavras chaves (metadados no servidor de conteúdo).

As atividades dos agentes coletores podem ser divididas em três etapas principais.

Na primeira etapa, um *agente reativo* recebe quais são os tópicos solicitados pelo professor e que existem no servidor, busca através dos elementos da meta-informação.

Na segunda etapa um *agente pró-ativo* tem em conta os tópicos do material novo, ele continua ligado ao professor informando quais são os tópicos que ingressarem.

E na terceira etapa, um *agente móvel* busca periodicamente as informações dos tópicos de conteúdos, generalizando para qualquer tipo de conteúdo em outros servidores, pois pode ser uma federação de servidores.

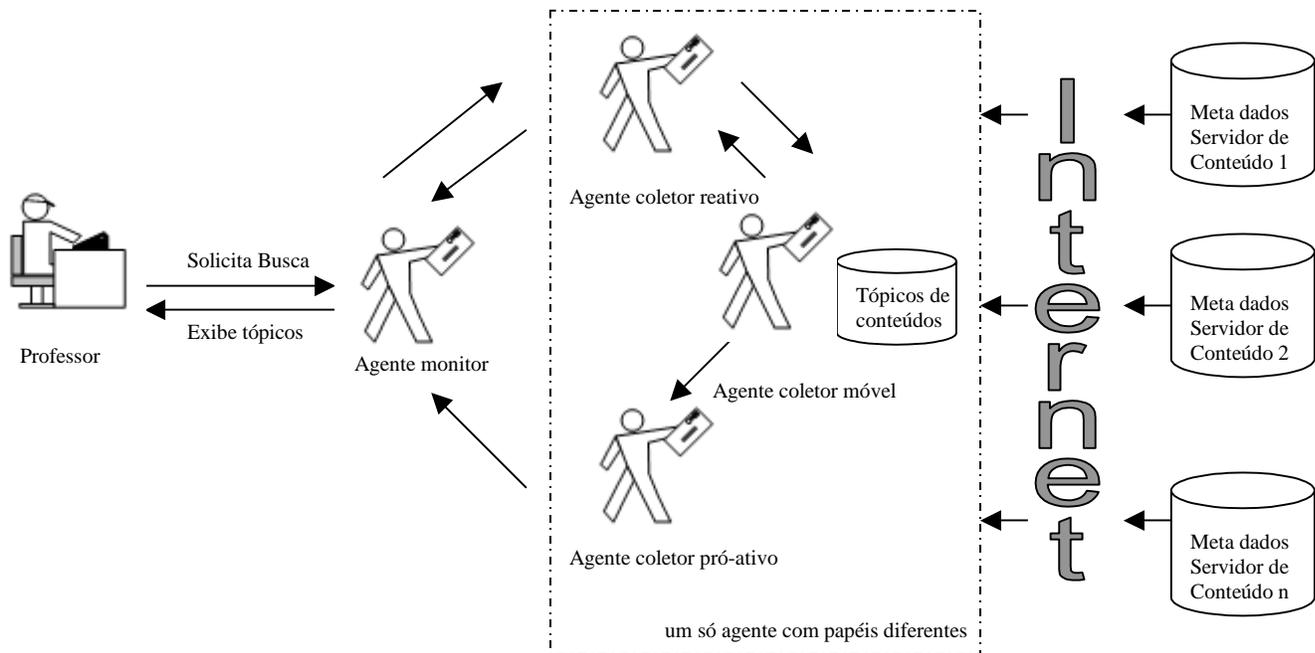


Figura 6. Passos de execução dos agentes para a realização da busca

O sistema deve implementar também, uma ferramenta que proporcione a adaptação dos cursos já existentes e dos novos cursos em um novo formato, o dos conteúdos em metadados dos conteúdos, para serem localizados e acessados pelos agentes. Esses metadados podem estar em arquiteturas baseadas em XML (*Extensible Markup Language*) [18] através do auxílio de Ontologias que suporta o processo de extração, detectando as informações interessantes ao domínio [17]. Desta forma, a arquitetura se torna flexível para atender a diferentes domínios do assunto.

6 Conclusão

Nos motores de busca, a pesquisa por informações na Web em geral é realizada por ferramentas de busca ou os conhecidos robôs de busca encontrados em sites especializados neste serviço, que armazenam as informações sobre páginas da Internet em um banco de dados. O usuário pode definir consultas através de palavras-chave, utilizando operadores do tipo *e, ou, não, perto, etc.* As consultas do usuário são realizadas sobre o banco de dados e as informações são recuperadas para o usuário informando a fonte de informação que contém a palavra-chave desejada.

Acreditamos que a utilização de agentes pode favorecer vários aspectos de interoperabilidade de produtos e conteúdos em educação à distância, estando centrada as atividades dos agentes nos seguintes focos:

- procura de informação que saibam como localizar os dados, que saibam como obter acesso ou mesmo negociar os modos de acesso, e que tenham informações sobre o cuidado da fonte;

- filtragem de informação que possuam modelos de seus usuários, propiciando uma busca mais efetiva e personalizada;
- facilitação e negociação de diferentes agentes interessados em obter informações ou conteúdos educacionais;
- monitoramento de dados em diferentes fontes de informação sejam elas servidoras de conteúdos educacionais ou não, apoiando de alguma forma a mobilidade dos aprendizes, professores ou projetistas.

O sistema apresentado, auxilia ao professor na busca de conteúdos educacionais, através da utilização de agentes que possibilitam o acesso, descrição, localização e uso dos conteúdos espalhados em servidores compatíveis com o padrão IMS, ajudando assim, ao professor, a melhorar e automatizar a precisão das pesquisas.

Referências Bibliográficas:

- [1] Nwana H. Software Agents: An Overview. Knowledge Engineering Review 1996; 11(3): 1-40.
- [2] Silva, V., Lucena, C. J. P. ContentNet: Um Framework para Interoperabilidade de Conteúdos Educacionais Utilizando o Padrão IMS. Dissertação para obtenção do título de Mestrado em Ciência em Informática. Puc-Rio, 2000.
- [3] Weiss G, editor. Multiagent Systems: modern approach to Distributed Artificial Intelligence. Cambridge: MIT Press; 1999.
- [4] Jennings N and Wooldridge M. "Agent-Oriented Software Engineering". In: J. Bradshaw (ed.), Handbook of Agent Technology, AAAI/MIT Press, 2000.
- [5] Franklin S and Graesser A. Is it an agent, or just a program?: A taxonomy for autonomous agents. ECAI'96 Workshop on Agents Theories, Architectures, and Languages: Intelligent Agents III, 21-36. Berlin: Springer.
- [6] Wooldridge M and Jennings N. Intelligent agents: Theory and Practice. The Knowledge Engineering Review 1995; 10(2): 115-152.
- [7] Bradshaw, J.M. An introduction to software agents In Bradshaw, J, M. Ed. Software Agents. Massachusetts: MIT Press, 1997.
- [8] Agent Technology – Green Paper, Version 1.0, Object Management Group, 01/08/2000.
- [9] Paker, Albert JAFMAS –A java based agent framework for multiagents systems. Cincinnati: Department of electrical & computer Engineering and computer Science University of Cincinnati, 1997. Doctoral thesis.
- [10] Minsky, M, A framework for representing knowledge. Mind design, 1981, pag. 95-128.
- [11] Sahuguet, A.; Azavant, F. Web Ecology: Recycling HTML pages as XML documents using W4F, In: Proceedings of the second international workshop on the web and databases, Philadelphia, USA, 1999.
- [12] IMS, 2000 “Instructional Management Systems Global Learning Consortium” disponível em <http://www.imsproject.org/> acessado em 10/04/2002.
- [13] Yahoo, disponível em <http://www.yahoo.com> acessado em 23/05/2002.

- [14] Callado, Luiz Fernando Martins, “WebBD-XML – Uma Arquitetura de um Sistema de Banco de Dados baseada em XML para extração de conteúdo da Web”. Dissertação para obtenção do título de Mestrado em Ciência em Informática. Puc-Rio, 2001.
- [15] IMS Learning Resources Metadata Specifications, disponível em <http://www.imsproject.org/metadata/mdinfov1pl.html> acessado em 14/04/2002.
- [16] Klusch, M. (Ed). Intelligent information agents: agent based discovery and management on the Internet. Springer-Verlag, 1999.
- [17] Pitts-Moultis, Natanya. Trabalhando com o RDF: Um Estudo de Caso XML Avançado. Capítulo 19, página 403. MAKRON Books.
- [18] Especificação XML Schema – Disponível em <http://www.w3.org/TR/xmlschema-0/> acessado em 15/05/2002.