BUSCA, RECUPERAÇÃO E VISUALIZAÇÃO DE RECURSOS ACADÊMICOS DISTRIBUÍDOS

Felipe N. Saab, Francisco A. da Silva, Silvio A. Carro

Faculdade de Informática - FIPP - Universidade do Oeste Paulista-UNOESTE - Presidente Prudente-SP. E-mail: silvio@unoeste.br

RESUMO

As universidades são referência quando o assunto é pesquisa e inovação. Tanta pesquisa gera muito material correlato (artigos e vídeos de experimentos por exemplo) e na maioria das vezes esse material não é compartilhado com a comunidade (seja qual for a natureza: acadêmica, científica ou pública) por falta de um meio que seja de fácil acesso e utilização. Vale lembrar que não apenas as pesquisas geram recursos acadêmicos, o ambiente de uma universidade oferece por si só um grande leque de atividades que geram recursos acadêmicos: aulas, iniciações científicas, simpósios, congressos, entre outros. O maior problema desse cenário é que tamanha riqueza de recursos não tem valor algum caso estes fiquem guardados. Como toda universidade dispõe de uma boa infra-estrutura em TI, a hospedagem e disponibilização de tais recursos na Internet é algo trivial, o problema é que deve-se conhecer onde os recursos estão para poder chegar até eles. É nesse contexto que este trabalho pretende desenrolar-se, permitir a busca destes recursos de um modo rápido e transparente ao usuário final.

Palavras-chave: infra-estrutura em TI –XML -recursos acadêmicos

SEARCH, RETRIEVAL AND VISUALIZATION OF DISTRIBUTED ACADEMIC RESOURCES

ABSTRACT

Universities are reference when the subject is research and innovation. So much research creates a lot of related material (papers and videos for example), mostly this material is not shared with the community (of any nature: academic, scientific or public) because there is not a means with easy access and use. Don't just researches create academic resources, the environment of a university by itself offer a wide range of activities that create academic resources: classes, scientific initiations, symposia, congress and others. The biggest problem of this scenario is that such a wealth of subjects doesn't have any value when they are saved. As every university has a good IT infrastructure, the hosting and providing of such resources on Internet is trivial, the problem is that should know where the resources are to get them. This work pretends to take place in this context, allow the search of such resources in a transparent way to the final user.

Keywords: IT infrastructure –XML - academic resources

INTRODUÇÃO

Como as universidades são grandes centros geradores e consumidores de recursos acadêmicos, a quantidade de recursos que são armazenados em sua própria infraestrutura de Tecnologia da Informação (TI) e até mesmo os que estão na Internet e são utilizados para realizar atividades relacionadas a pesquisa e ensino é muito grande.

Tais recursos poderiam ser utilizados com mais frequência para fins de pesquisa e ensino por outras instituições e/ou pessoas. Porém, a distribuição de tais recursos não é algo costumeiro nos dias de hoje. Um dos motivos é que não existe uma ferramenta de fácil acesso e/ou utilização difundida entre as universidades/pessoas que desejam compartilhar tais recursos.

Nesse contexto foi criado o trabalho MetaAcad [Caçula 2007] cujo objetivo é facilitar a interação de qualquer pessoa aos recursos disponíveis em um meio acadêmico, segundo Coelho (2010).

O MetaAcad propõe 3 módulos, sendo eles:

- o módulo descritor que anexa metadados modelados em XML (*Extensible Markup Language*) ao recurso e o envia para um serviço Web;
- o módulo gerenciador que é responsável por organizar e distribuir os recursos;
- o módulo de pesquisa que é responsável por recuperar tais recursos e disponibilizá-los para o usuário final.

Tais módulos foram desenvolvidos em trabalhos paralelos: o descritor MetaAcad Plus [Rebonatti 2010], e o Gerenciador e Recuperador de Recursos [Coelho 2010].

O problema dos módulos desenvolvidos é que eles funcionam de tal maneira que os meios acadêmicos estão sempre isolados.

No decorrer deste artigo é demonstrada como toda a estrutura já pronta do modelo MetaAcad pode ser reaproveitada e integrada com novas funcionalidades para se tornar um sistema distribuído que permita vários meios acadêmicos compartilharem e recuperarem seus recursos por meio de uma só ferramenta.

No decorrer do texto deste artigo é descrito inicialmente o MetaAcad Plus, que é um aplicativo desenvolvido para auxiliar na descrição de um recurso acadêmico multimídia. Após, é descrito o módulo gerenciador e recuperador de recursos, responsável por efetuar a organização, pesquisa e distribuição de recursos acadêmicos. Na sequência é descrito o módulo Gerenciador de Servidores que cuida da intercomunicação entre servidores para compartilhar recursos acadêmicos de forma distribuída por meio da Internet. A seção posterior trata da integração dos módulos Descritor e Gerenciador de Recursos. Por fim, encontram-se as conclusões e considerações finais do trabalho.

METAACAD PLUS

Tão importante quanto ter informação é conseguir encontrar o que se procura e checar se esta informação é de fonte fidedigna. Há muito tempo as pessoas não toleram passar horas procurando aquilo que buscam na internet ou mesmo em um site específico.

No mundo acadêmico não é diferente. A necessidade de uma ferramenta que ofereça o compartilhamento e facilidade na busca das informações produzidas é de grande importância na evolução das instituições [Rebonatti 2010].

O MetaAcad Plus é um aplicativo que visa auxiliar na descrição de um recurso acadêmico

multimídia, o qual envolve imagens, vídeos, áudio, apresentações e produções bibliográficas.

O sistema apresenta um conjunto de metadados apropriados para representar com detalhes cada recurso. Cada tipo de recurso (imagens, vídeos e outros) possui um conjunto de metadados específicos.

O conjunto de metadados é baseado no modelo MetaAcad [Caçula 2007] e descrito no formato XML. A Figura 1 mostra os metadados utilizados na descrição de dois recursos, imagem e vídeo.

```
<metaacad mime="JPEG">
 <recurso>
  <dt__criacao>21.04.2010</dt__criacao>
  <descricao>Palestra de empreendedorismo</descricao>
  <idioma>Portugues</idioma>
  <local__criacao>FIPP</local__criacao>
  <data__descricao>10.06.2010</data__descricao>
  <origem>www.unoeste.br/fipp</origem>
  <direitos>FIPP</direitos>
  <palavras__chave>palestra, fipp,
empreendedorismo</palavras__chave>
 </recurso>
 <criador>
  <nome></nome>
  <email></email>
  <modalidade></modalidade>
 </criador>
 <multimidia tipo="Imagem">
  <personagem>
   <nome__personagem></nome__personagem>
  </personagem>
  <imagem>
   <dimensao>700x700</dimensao>
   <formato__cena>JPEG</formato__cena>
  </imagem>
 </multimidia>
</metaacad>
```

```
<metaacad mime="MPEG">
 <recurso>
  <dt__criacao>20.05.2010</dt__criacao>
  <descricao>Aula de matematica</descricao>
  <idioma>portugues</idioma>
  <local__criacao>FIPP</local__criacao>
  <data__descricao>09.06.2010</data__descricao>
  <origem>www.unoeste.br</origem>
  <direitos>FIPP</direitos>
  <palavras__chave>aula, matematica</palavras__chave>
 </recurso>
 <criador>
  <nome></nome>
  <email></email>
  <modalidade></modalidade>
 </criador>
 <multimidia tipo="Video">
  <personagem>
   <nome__personagem></nome__personagem>
  </personagem>
  <video>
   <dimensao>800x600</dimensao>
   <duracao>5 minutos</duracao>
  </video>
 </multimidia>
</metaacad>
```

Figura 1. Metadados utilizados na descrição de dois recursos dos tipos imagem e vídeo

O sistema permite que a descrição de cada recurso seja anexada a ele, ou seja, o XML contendo a descrição é incorporado no corpo do arquivo. Desta maneira, o recurso leva sua descrição automaticamente para qualquer lugar que seja transferido.

Os recursos, juntamente com sua descrição são armazenados em um servidor multimídia e posteriormente podem ser recuperados por outros sistemas ou manualmente via mecanismos de busca apropriados [Rebonnati 2010].

O sistema é executado tanto em ambiente Web como em ambiente Desktop, pois foi escolhida a tecnologia JavaFX para o desenvolvimento dele. A Figura 2 mostra a tela utilizada para descrever um recurso.



Figura 2. Tela do MetaAcad Plus em execução no ambiente Desktop mostrando a descrição de uma imagem

GERENCIADOR E RECUPERADOR DE RECURSOS

Este módulo objetiva a especificação e implementação de um modelo gerenciador responsável por efetuar a organização, pesquisa e distribuição de recursos acadêmicos gerados em uma instituição e documentados em XML. O sistema utiliza a Internet como repositório de recursos e uma aplicação desenvolvida utilizando a tecnologia Java para a recuperação desses recursos (JSP, Servlets e Web Services) [Coelho 2010].

Tal trabalho tem como principais funções a publicação de novos recursos (já descritos pelo descritor MetaAcad Plus) e a busca de tais recursos a partir de palavras chave e/ou parâmetros específicos de uma mídia (imagem, áudio, documento .pdf, entre outros).

A publicação pode ser realizada a partir de dois caminhos: envio do recurso encapsulado em uma mensagem SOAP (*Simple Object Access Protocol*) para o serviço Web ou por meio da interface de publicação de recursos no aplicativo Web.

Todo recurso publicado no servidor tem seus metadados armazenados em um banco de dados relacional, o que permite a rápida busca e filtragem dos recursos a partir dos parâmetros de busca. A Figura 3 mostra o diagrama de Classes de Objetos referente aos metadados de recursos acadêmicos utilizados para armazenar os dados desses recursos no banco de dados.

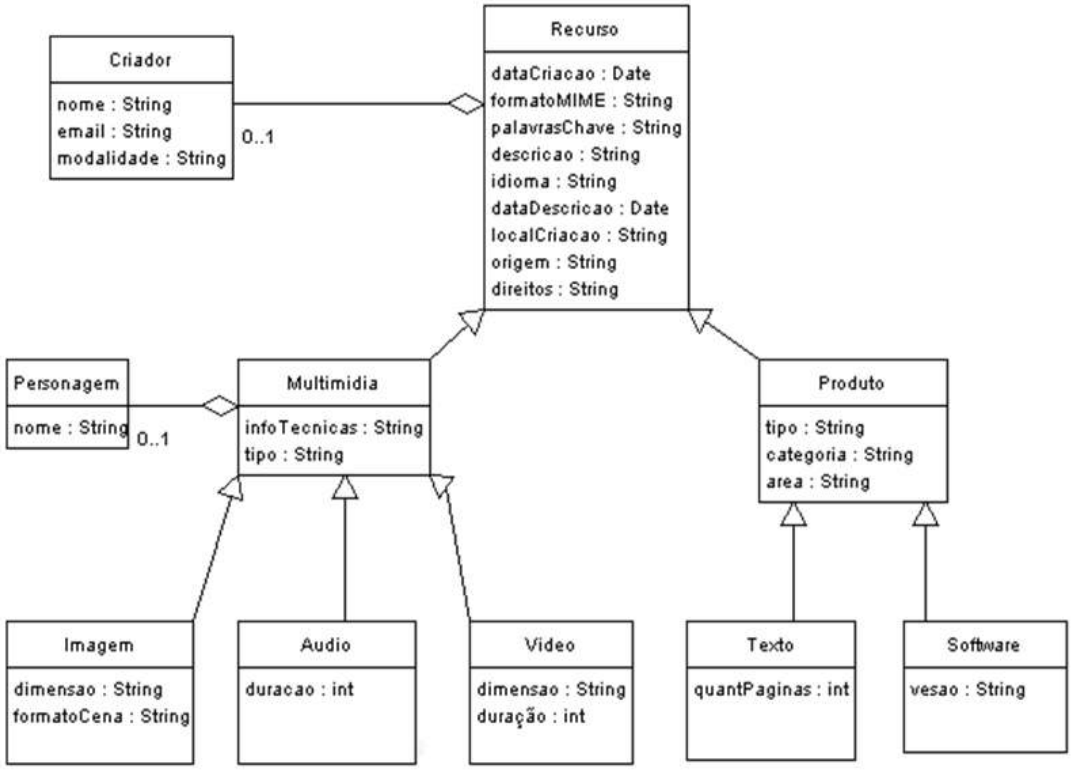

Figura 3. Diagrama de Classes de Objetos referentes aos Metadados de Recursos Acadêmicos
Fonte: [Caçula 2007].

A busca por palavras chave é ampla, são feitas buscas em quaisquer campos do conjunto de metadados do recurso. Uma busca pela palavra chave '*auditoria*' irá recuperar registros cuja palavra chave apareça em qualquer atributo do tipo *String*.

Já a busca avançada permite realizar buscas em campos específicos, como por exemplo, na dimensão de uma imagem ou o idioma do recurso.

GERENCIADOR DE SERVIDORES

A natureza da Internet propicia uma boa oportunidade para o compartilhamento de recursos acadêmicos que sejam previamente descritos e publicados em servidores na rede.

O maior problema reside no fato do escopo do módulo Gerenciador e Recuperador de Recursos [Coelho 2010] se restringir a apenas um servidor, ou seja, não há recuperação de recursos que estão publicados em outros servidores, a busca funciona localmente.

Com base nesse problema foi desenvolvido um módulo Gerenciador de Servidores que cuida da intercomunicação entre servidores de entidades que visam compartilhar recursos acadêmicos de forma distribuída através da Internet. Tal módulo mantém um grupo de servidores se comunicando no estilo P2P – todos se comunicando com todos.

O módulo Gerenciador de Servidores implementa as seguintes operações:

- Inclusão de servidores no grupo (Figura 4): um dos servidores que fazem parte do grupo efetua a inclusão de um novo servidor através do seu endereço IP. Durante o processo de inclusão todos os servidores do grupo passam a conhecer o novo servidor e ele também conhece todos os servidores já existentes no grupo;
- Auto exclusão de um servidor do grupo (Figura 5): quando um servidor não quer mais participar do grupo ele

efetua a sua auto exclusão avisando todos os outros servidores que está se retirando do grupo e depois passa a desconhecer os servidores do grupo limpando seu banco de dados;

- Registro e posterior execução de pendências (Figura 6): caso durante a inserção de um novo servidor exista algum servidor do grupo que não esteja online, o servidor que está realizando a inserção vai guardar uma pendência de que aquele servidor não conhece o que está sendo inserido agora. No momento em que os servidores conseguirem se comunicar novamente o que estava offline irá ser avisado de que houve a inserção de um novo servidor e todos os servidores conhecerão todos novamente. Esta funcionalidade está disponível para o processo de auto exclusão também, caso algum servidor esteja offline durante o processo de exclusão será escolhido outro servidor online para guardar a pendência e posteriormente avisar que um servidor saiu do grupo.

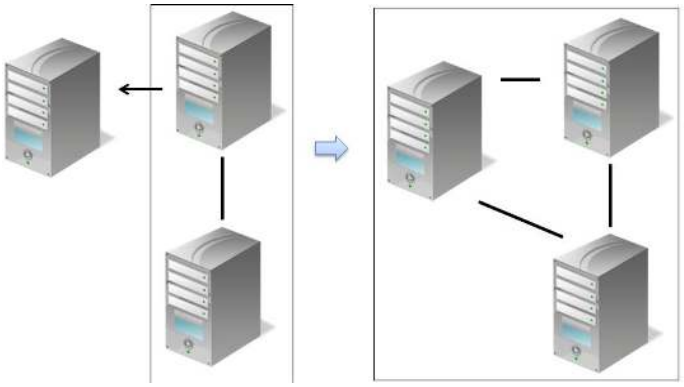

Figura 4. Inclusão de um servidor no grupo

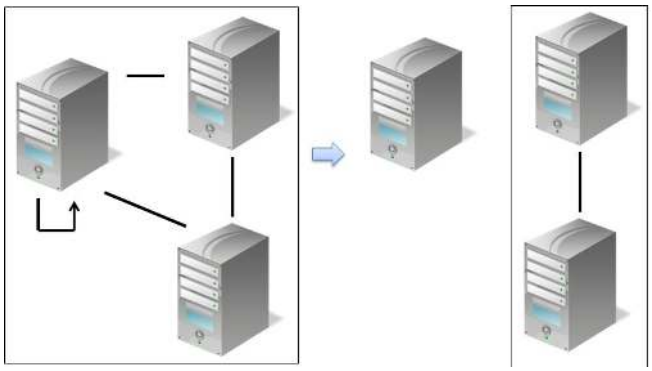

Figura 5. Autoexclusão de um servidor do grupo

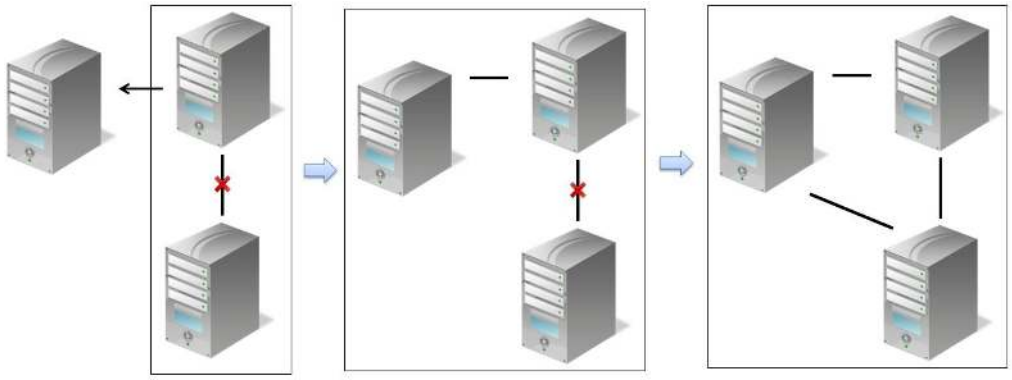

Figura 6. Registro de pendência para o caso de inclusão

Toda a comunição entre os servidores é realizada por meio da troca de mensagens através da Internet utilizando Web Services [Topley 2003] com base nas tecnologias WSDL (*Web Service Definition Language*) [Christensen e Curbera 2010] e SOAP [Gudgin, Hadley e Mendelsohn 2010].

TODOS OS MÓDULOS TRABALHANDO EM CONJUNTO

Uma vez que os módulos do modelo MetaAcad (Descritor e Gerenciador de Recursos) e o Gerenciador de Servidores estão desenvolvidos, existe um conjunto de aplicações que podem se integrar para atingir um objetivo em comum: facilitar a interação das pessoas com os recursos acadêmicos.

O primero passo foi a integração dos módulos Descritor e Gerenciador de Recursos. Entrando mais em detalhes, a integração deles é referente ao requisito '*publicação*' de um recurso acadêmico, ou seja, uma vez que o usuário final termina de descrever o recurso, ao invés de apenas anexar a descrição no recurso e depois publicá-la através do módulo Gerenciador, a publicação pode ser realizada de forma direta. Após a descrição, o recurso é publicado no servidor para o qual ele está configurado.

A Figura 7 mostra o fluxo das informações para a publicação de um recurso (incluindo sua descrição).



Figura 7. Anexação da descrição no recurso e publicação através da Internet com apenas um clique

Como o módulo Gerenciador utiliza Web Services para receber os recursos, as publicações podem ser remotas, ou seja, qualquer pessoa autorizada poderia publicar recursos acadêmicos em um servidor que esteja executando o módulo Gerenciador.

Com o modelo MetaAcad implementado e funcionando (módulo Descritor e módulo Gerenciador de Recursos), a idéia de expandí-lo ainda mais é possível através do módulo Gerenciador de Servidores.

O módulo Gerenciador de Servidores é a base que permite o módulo Gerenciador e Recuperador de Recursos local de um servidor se comunicar com os módulos de outros servidores do grupo, ou seja, extende a funcionalidade de busca e recuperação de recursos ao escopo distribuído.

A Figura 8 demonstra como funciona a busca de recursos em um escopo distribuído, o usuário faz a requisição a um servidor e o módulo Gerenciador de Servidores se encarrega de enviar a busca para todos os servidores do grupo. Ao receber as requisições, os servidores fazem uma busca local e retornam seus resultados para o servidor que encaminhou as requisições. A partir de então o servidor apenas formata os resultados e exibe uma visualização amigável para o usuário final, que provavelmente nem sabe que a busca foi realizada de forma distribuída.

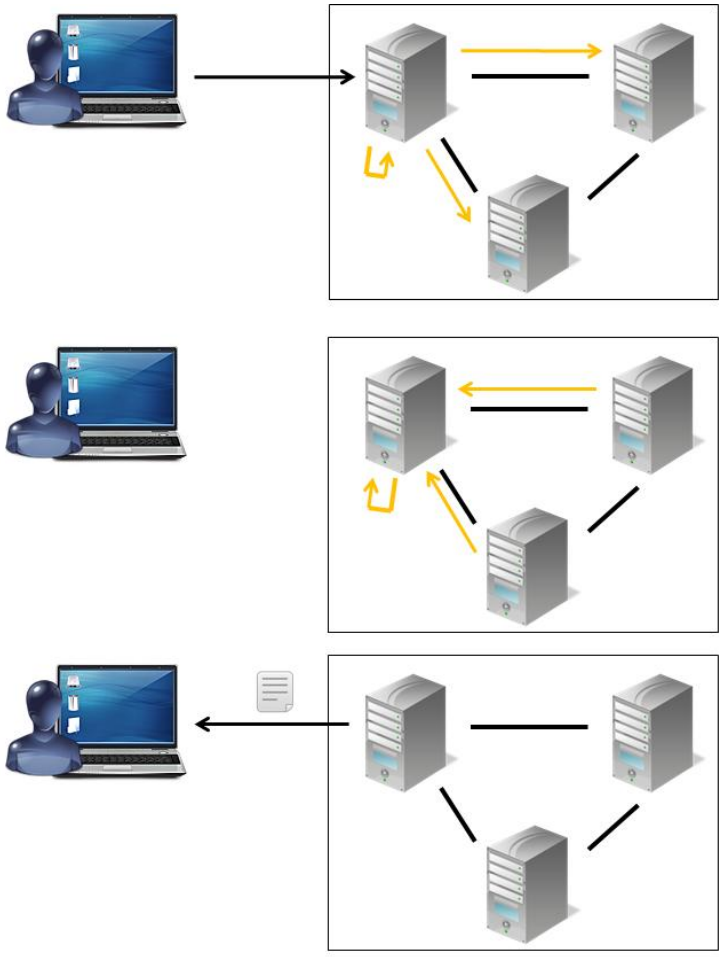

Figura 8. Busca distribuída possibilitada pelo módulo Gerenciador de Servidores

CONCLUSÕES E CONSIDERAÇÕES FINAIS

Com todos os módulos básicos do modelo MetaAcad implementados e integrados torna-se possível o compartilhamento de recursos acadêmicos em uma ou mais instituições.

Com a integração do módulo Gerenciador de Servidores aos dois outros módulos, esta funcionalidade se tornou realidade também em um ambiente distribuído, ou seja, intituições de vários lugares diferentes podem se conectar facilmente através da Internet para compartilhar seus recursos acadêmicos e manter uma rede que tende a crescer cada vez mais com o passar do tempo.

Contudo, existem muitas possibilidades de aperfeiçoar este trabalho que devem ser levadas em consideração futuramente. São elas:

- Indexação local dos recursos distribuídos para evitar muitas chamadas aos Web services de busca;
- Bloqueio da publicação de recursos duplicados. Tal funcionalidade não pode se restringir apenas à comparação dos metadados, deve-se utilizar recursos mais avançados para a compação dos diversos tipos de recursos;
- Filtragem dos recursos na recuperação. Uma vez que o impedimento de publicação de recursos duplicados é executado no escopo local, uma busca distribuída pode ocasionalmente resultar em recursos duplicados;
- Classificação e ordenação dos recursos resultantes de uma busca. Quais são os mais relevantes e, portanto, devem ser exibidos primeiro é uma questão que deve ter sua solução muito bem planejada. Uma opção seria extender os metadados do MetaAcad para possibilitar sua classificação.

Muitas outras idéias podem surgir no decorrer das próximas modificações deste aplicativo que tende a melhorar e ajudar muitas pessoas a encontrar recursos acadêmicos para aproximá-las do aprendizado.

REFERÊNCIAS

CAÇULA, M. E. F. MetaAcad – Descritor: Uma aplicação para a descrição de recursos. Universidade do Oeste Paulista, Presidente Prudente – SP. 2007.

CHRISTENSEN E., CURBERA F. et. al. Web Services Description Language (WSDL) 1.1. World Wide Web Consortium. Acessado em 15 de Mar. de 2010. Disponível em http://www.w3.org/TR/wsdl.

COELHO, S. R. N. Distribuição, recuperação e autoria de recursos acadêmicas com base no modelo MetaAcad. Universidade do Oeste Paulista, Presidente Prudente – SP. 2010.

GUDGIN M., HADLEY M., MENDELSOHN N., et. al. SOAP Version 1.2 Part1: Messaging Framework. World Wide Web Consortium. Acessado em 27 de Abr. de 2010. Disponível em: http://www.w3.org/TR/soap12-part1/.

REBONNATI, M. F. MetaAcad Plus: Uma extensão do descritor MetaAcad para a descrição de objetos multimídia na área acadêmica. Universidade do Oeste Paulista, Presidente Prudente – SP. 2010.

TOPLEY K. Java Web Services in a Nutshell. O'Really Media, 2003.