

Sistema Colaborativo Baseado em Modelos Semânticos para Compartilhamento de Documentos

Renan Manola, Renata S.S. Guizzardi, Roberta Lima Gomes

Universidade Federal do Espírito Santo (UFES)
Av. Fernando Ferrari, S/N – 29060-970 – Vitória/ES

rmanola@gmail.com, {rguizzardi, rgomes}@inf.ufes.br

***Abstract.** This paper describes Biblioref, a collaborative system to support document sharing. In this system, document classification and sharing is based on the development of user's individual taxonomies, which are mapped into a reference taxonomy. The main benefits of this system can be described as: a) promoting collaboration among users through mutual access to each other's documents; b) granting user autonomy in knowledge organization, since documents are classified according to individual points of view instead of a centralized model; and c) providing mechanisms to relate the different user's classification schemes, allowing to find potential collaborators, inferred from these relations.*

***Resumo.** Este artigo descreve Biblioref, um sistema colaborativo para compartilhamento de documentos. A organização e busca de documentos se baseiam na construção de taxonomias individuais que são mapeadas em uma taxonomia de referência. Os principais benefícios do sistema são: a) promover a colaboração entre usuários, permitindo acesso mútuo a documentos; b) fornecer autonomia ao usuário na organização de conhecimento, já que os documentos são classificados a partir de visões individuais ao invés de um modelo central; e c) prover mecanismos de relacionamento entre as múltiplas categorizações, permitindo inferir colaboradores em potencial a partir de tais relacionamentos.*

1. Introdução

O crescente desenvolvimento das tecnologias da informação inaugurou novas maneiras de lidar com o trabalho, bem como estruturar empresas e outras organizações. Algumas dessas organizações se tornaram parcial ou completamente virtuais, seus processos de trabalho ganharam natureza mais dinâmica e distribuída, e suas estruturas, anteriormente estáticas e altamente hierarquizadas, passaram a ser mais adaptáveis e flexíveis.

Nesse contexto, o suporte ao trabalho colaborativo e à gestão do conhecimento torna-se ainda mais importante e chega a ser considerado por alguns autores [Nonaka e Takeuchi 1995; Fischer e Ostwald 2001; Guizzardi 2006] como vital para viabilizar o funcionamento das organizações modernas. Em suma, tal trabalho refere-se a uma área de pesquisa que estuda como pessoas que se localizam *geograficamente distribuídas* podem se estruturar e interagir para efetuar *atividades* que levam a obter um *objetivo comum* [Gomes 2006]. Já a gestão do conhecimento trata da *criação, integração e uso do conhecimento*, com a meta principal de *melhorar o desempenho* de indivíduos e

organizações [Fischer and Ostwald 2001]. Dessa forma, essas duas áreas se complementam para garantir que indivíduos e organizações se mantenham competitivos na sociedade da informação.

Atualmente, do que se conhece de sistemas de suporte à colaboração e à gestão do conhecimento, percebe-se que a tendência é prover portais e repositórios centrais, que assumem vocabulários, linguagens e esquemas de classificação centralizados [Liao, 2002]. Conseqüentemente, a falta de confiança e motivação dos colaboradores leva à insatisfação [Pumajera *et al* 2003; Merali e Davies, 2001]. Em outras palavras, usuários resistem a trocar conhecimento, já que eles não sabem quem vai capturá-lo e o que será feito com ele. Conseqüentemente, esses sistemas são frequentemente deixados de lado, e as pessoas continuam colaborando e trocando conhecimento usando processos já difundidos (e.g. questionando colegas do seu ciclo de confiança, enviando email a pessoas reconhecidas como experts em determinado assunto etc.) [Pumajera *et al* 2003; Merali e Davies, 2001].

Este artigo apresenta Biblioref, um sistema colaborativo para compartilhamento de documentos. Tal sistema tem três funcionalidades (benefícios) principais:

- permitir que os usuários possam ter documentos (em formato PDF) armazenados em um sistema de gerenciamento de conteúdo (*Content Management System*) baseado na *World Wide Web*;
- permitir que usuários tenham uma visão personalizada de como o sistema é organizado e de como seus documentos são armazenados, obtendo, portanto, mais autonomia a partir da criação de modelos individuais de classificação, ao invés de organizar e visualizar documentos unicamente com base em um modelo central;
- permitir mecanismos de mapeamento entre os esquemas de classificação dos diversos usuários, possibilitando que colaboradores em potencial sejam inferidos a partir de tais relacionamentos. O estabelecimento de relações entre esses esquemas é transparente aos usuários finais, oferecendo-lhes a opção de querer ou não fazer uso desses mapeamentos.

Além dessa introdução, este artigo conta com mais quatro seções, estruturadas da seguinte forma: a seção 2 descreve os principais aspectos do sistema Biblioref; a seção 3 apresenta alguns trabalhos relacionados; a seção 4 discute o uso de modelos semânticos como base para sistemas colaborativos de gestão de conhecimento; e, finalmente, a seção 4 traz as considerações finais deste trabalho.

2. Biblioref: um Sistema Semântico para Gerenciamento de Referências Bibliográficas

Biblioref é um sistema baseado na *Web* que oferece suporte à gestão de conteúdo a partir de documentos que compõem referências bibliográficas mantidas por profissionais da área acadêmica. Esse conteúdo compõe-se, principalmente, de artigos científicos, dissertações de mestrado, teses de doutorado, relatórios técnicos, entre outras publicações.

Biblioref é baseado em um sistema de gerenciamento de conteúdo (*Content Management System* – CMS) existente, denominado *Drupal*¹. Tal escolha foi baseada em um estudo aprofundado, abrangendo diversos CMSs. Ao fim desse estudo, concluiu-se que o *Drupal* era a melhor opção, em especial por sua tendência de crescimento e adoção cada vez maior, além de amplo suporte para desenvolvedor.

2.1. Organização e Busca de Documentos baseada em Modelos Semânticos

O sistema permite que cada usuário crie uma ou mais taxonomias para descrever as áreas de conhecimento abrangidas pelos documentos que coleciona. Do ponto de vista do usuário, essa taxonomia pode ser simplesmente considerada uma estrutura de arquivos que se assemelha às estruturas utilizadas para organizar conteúdo em sistemas operacionais tradicionais. Dessa forma, ao invés de assumir que o usuário comum é capaz de criar estruturas complexas tais como Ontologias de Domínio, prefere-se adotar uma abordagem mais direta e clara, já conhecida por ele, para facilitar a utilização do sistema. Ao mesmo tempo, tal classificação individual serve de base para a visualização, organização e busca de documentos.

A busca de documentos é viabilizada pelo mapeamento de, por exemplo, uma taxonomia individual sobre um determinado domínio em uma taxonomia de referência que o sistema provê para esse mesmo domínio.

No sistema Biblioref existem basicamente três tipos de relação:

1º.Documento ↔ Termo da taxonomia do usuário

2º.Termo da taxonomia do usuário ↔ Termo da taxonomia de referência

3º.Documento ↔ Termo da taxonomia de referência

A primeira relação é efetuada pelo usuário, que especifica manualmente a que termo de sua taxonomia um documento cadastrado será associado. A segunda relação pode ser estabelecida automaticamente (ou, no pior caso, semi-automaticamente), por meio do algoritmo explicitado na Figura 1. A terceira relação também é inferida automaticamente; a princípio, o sistema usa o algoritmo descrito na Figura 1 para inferir uma relação entre o documento e um termo da taxonomia de referência; no caso de obter resultado nulo, o sistema se apóia na segunda relação, ou seja, o documento é ligado ao termo da taxonomia de referência associado ao termo da taxonomia do usuário.

Como mostra a Figura 1, a lógica encarregada de inferir essas relações é baseada em um sistema de *ranqueamento*. Nessa figura, considera-se um termo **T** sendo cadastrado pelo usuário, sendo que **T** tem sinônimos **S**. Busca-se, aqui, relacionar o termo **T** ao termo **T_ref** da taxonomia de referência, sendo que **T_ref** tem sinônimos **S_ref**. No algoritmo, retornos NULL ou 0 (zero) indicam, respectivamente, que o sistema não consegue obter um ranqueamento ou não consegue desempatar ocorrências. Nesse caso, o usuário é solicitado a indicar explicitamente como a relação deve ser estabelecida. Essa decisão de deixar que o usuário interfira somente em última instância reflete o compromisso de tentar fazer um sistema que seja o mais automatizado possível, reduzindo, assim, o esforço despendido pelo usuário na classificação de

¹ <http://api.drupal.org>

documentos. No entanto, caso o sistema não consiga ter o comportamento esperado, o usuário deve ser consultado, garantindo-se que o resultado final não seja comprometido.

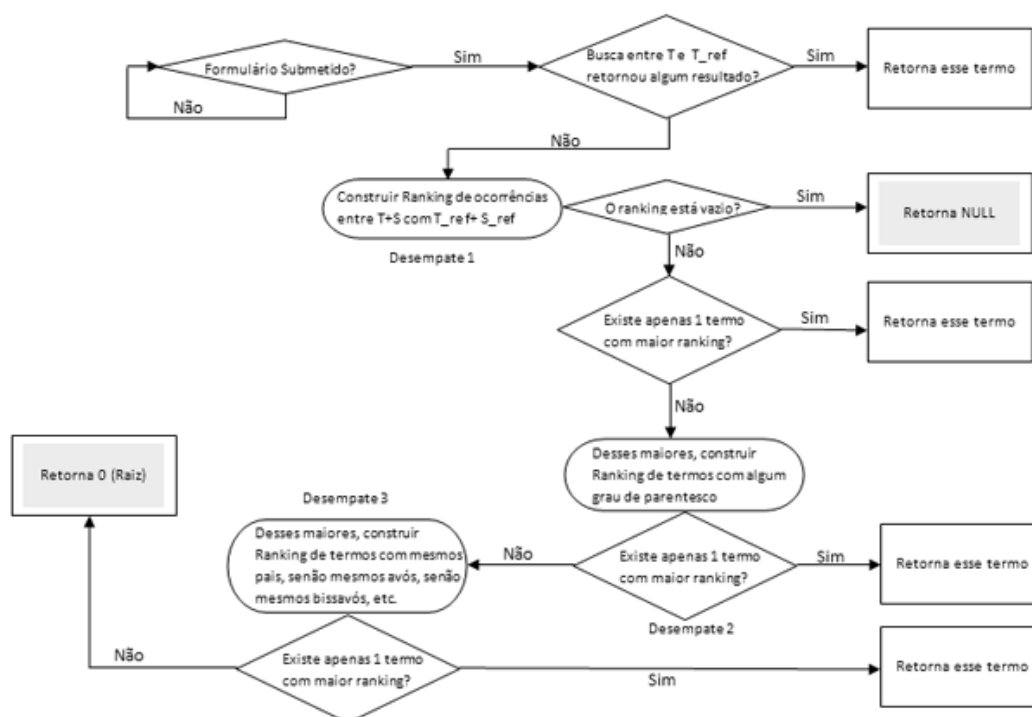


Figura 1. Algoritmo usado no estabelecimento automático da relação do tipo: “Termo da taxonomia do usuário” – “Termo da taxonomia de referência”.

Observe que um possível cenário de ocorrência onde o sistema retornaria NULL pelo **Desempate 1** (ver Figura 1), seria quando o sistema não pôde detectar qualquer semelhança entre um termo sendo cadastrado e os pré-existentes na taxonomia de referência. Isso poderia ocorrer pelo fato do termo sendo cadastrado fugir do escopo do domínio de classificação pressuposto, ou uma taxonomia de referência incompleta, talvez mal elaborada. Essa última questão pode ser evitada caso o modelo de referência seja elaborado a partir de uma ontologia de domínio consistente (ver discussão proposta na seção 4).

Para deixar ainda mais claras as três relações supracitadas, a Figura 2 mostra um exemplo visual de um cenário de melhor caso nos três tipos de relações existentes no sistema. Note que o documento **DOC** foi associado pelo usuário ao termo **UTA** de sua taxonomia pessoal (1ª relação). Note também que o sistema associou o termo **UTA** da taxonomia do usuário ao termo **REF** na taxonomia de referência (2ª relação). No entanto, o documento **DOC** relacionado ao termo **RRF** da taxonomia de referência (3ª relação). Isso ocorre porque, nesse caso, o algoritmo da Figura 1 pôde inferir automaticamente a terceira relação (Documento – Taxonomia de Referência). Se esse algoritmo não conseguisse determinar uma relação entre **DOC** e um termo da taxonomia de referência, a tomada de decisão automática do sistema seria de atrelar **DOC** com **REF** (e não **RRF**), ou seja, ao termo correspondente a **UTA** na taxonomia do usuário.

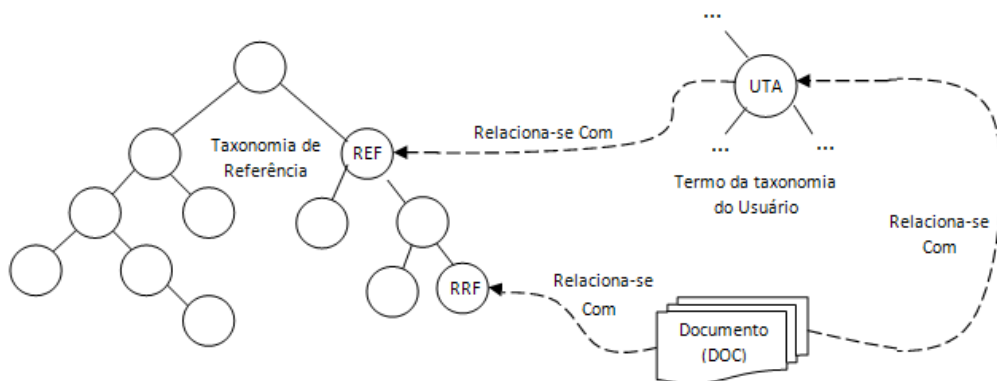


Figura 2. Relações entre taxonomias individuais, taxonomia de referência e documentos.

2.2. Funcionamento Básico do Sistema

A tela inicial do sistema é exibida na Figura 3.

Biblioref
Semantic Bibliographic Management System

Navigation

- Getting Started
- Recent posts
- 000WebHost

User login

Username:

Password:

Login

- Create new account
- Request new password

Welcome to Biblioref!

Here you'll be able to store, recover and share your references with other people, at the beginning you may find important to read the "Getting Started" guide located in the menu at the left of the screen. Never mind bother yourself the concerns of carrying with you all your bibliographic data, here you'll have things organized just like it is at your computer at home!

Recent posts

Type	Post	Author	Replies	Last updated
Biblioref	A survey of peer-to-peer content distribution technologies	rmanola	0	2 min 6 sec ago
Biblioref	Defining and exploring a grid system ontology	rmanola	0	4 min 48 sec ago
Biblioref	Wiki content templating	rmanola	0	7 min 59 sec ago
Biblioref	Supporting creativity in distributed scientific communities	rmanola	0	11 min 28 sec ago
Biblioref	Collaboration transparency in the DISCIPLINE framework	rmanola	0	31 min 36 sec ago

Free Hosting Service: 000WebHost

Figura 3. Tela inicial do sistema Biblioref

Inicialmente, é exibida uma tela de boas-vindas e, logo abaixo, um quadro contendo os documentos que foram recentemente inseridos. Para cada documento, é exibido o título e o usuário que o inseriu. Assim, oferece-se uma noção de atualização constante, mesmo que posteriormente o usuário efetue *login* e passe a ter sua visão particular do sistema. Um usuário que não possui conta pode clicar nos *links* e acessar os conteúdos, uma vez que o sistema é feito com o objetivo de disponibilizar informação a todos.

A primeira ação de um novo usuário é a de criar seu esquema de classificação individual, cadastrando novos termos, de forma que o sistema fique o mais parecido com sua organização de documentos local possível. Nesse processo, a inserção de sinônimos para cada termo é importante, pois como ressaltado na seção 2.1, o sistema faz uso desses metadados, ao tentar inferir automaticamente relações entre termos do usuário e termos da taxonomia de referência. A partir disso, o sistema também é capaz de inferir relações entre um documento sendo cadastrado e um termo da taxonomia de referência.

Após criar pelo menos parte de sua taxonomia individual, o usuário está pronto para organizar seus documentos. A Figura 4 exibe o formulário de cadastro de referências bibliográficas. Na criação de conteúdo, o primeiro campo a ser especificado é o de **My Taxonomy**, em que o usuário deve indicar em qual termo de sua taxonomia esse documento deve ser relacionado. Após o cadastramento do título e do *abstract*, o usuário deve colocar as palavras-chave que caracterizam o documento; essas palavras-chave podem originar-se de campos específicos de certos tipos de documento² ou ser definidas pelo próprio usuário.

Na Figura 5, pode-se verificar como documentos são listados, quando o usuário faz uma busca relacionada a termos de sua taxonomia.

Home » Create content

Submit Biblioref

Fill the fields above with information about the document.

My Taxonomy: *

Selecione a taxonomia para o documento que está sendo cadastrado

Title: *

Abstract:

Insira o Abstract do artigo, normalmente encontra-se no começo do mesmo.

Keywords: *

Insira as palavras-chave desse documento separadas por vírgulas

Authors: *

Insira os nomes dos autores, separados por vírgulas.

Upload de Referência

Changes made to the attachments are not permanent until you save this post.

Attach new file:

Allowed extensions: pdf

Figura 4. Tela de cadastro de uma nova referência bibliográfica.

Biblioref
Semantic Bibliographic Management System

rmanola

- My account
- Add Reference
- Edit My Taxonomy
- Getting Started
- Recent posts
- Log out

Knowledge Management

A theoretical framework for knowledge management implementation
 Posted July 28th, 2008 by rmanola Knowledge Management Strategies **2**

Views:
 Read (swf) | Read (pdf) | Download

Wiki content templating
 Posted July 27th, 2008 by rmanola Knowledge Management Wiki **3**

Views:
 Read (swf) | Read (pdf) | Download

Content Browser

- Artificial Intelligence
- Collaborative Web Systems
- Creative Support
- Knowledge Management
- Group Awareness
- Peer-to-Peer Systems

Free Hosting Service: 000Webhost

Figura 5. Tela que lista os documentos na navegação por termos, mostrando, em (1) a taxonomia de usuário e em (2) e (3) documentos relativos ao termo selecionado nessa taxonomia.

A visualização dos documentos da Figura 5 corresponde a uma pré-visualização, em que são exibidos apenas o título, os termos relacionados e as opções de leitura de cada documento. Nas opções de leitura, o usuário pode: a) ler o documento no browser fazendo uso de um plug-in do Flash ; b) ler o documento usando um plug-in de leitura de documentos em PDF que é embutido no browser; ou c) efetuar um download. Ao

² Em geral, artigos acadêmicos (tais como artigos científicos, dissertações e teses) possuem campos específicos onde são expostas palavras-chaves do mesmo.

clicar no título da pré-visualização de um dos documentos listados (Figura 5 (2) e (3)), o usuário é redirecionado à tela de exibição completa do documento.

A Figura 6 mostra um exemplo da tela de visualização completa. Essa visualização mostra mais detalhes do documento cadastrado, tais como as árvores de classificação dos termos associados (destacado por um retângulo). Com essa visualização, é possível saber em qual contexto um documento se encontra na classificação própria de um usuário, bem como no modelo de referência. A visualização do *abstract* e das palavras-chave é outra estratégia de prover informações relevantes para o usuário, dando suporte à decisão de acessar ou não tal referência.

The screenshot displays the Biblioref Semantic Bibliographic Management System interface. The main content area shows a document titled "Wiki content templating" posted on July 27th, 2008, by user "rmanola". The document includes an abstract, keywords ("microformats, semantic web, template, wiki engine"), and authors ("Angelo Di Iorio, Fabio Vitali, Stefano Zacchiroli"). Below the document content, a box titled "Aggregated Taxonomies" is highlighted with a red rectangle. This box contains two taxonomies: "rmanola's Taxonomy" with sub-terms "Collaborative Web Systems" and "Knowledge Management"; and "Reference Taxonomy" with sub-terms "Network Collaboration", "CMS", and "Wiki". The interface also features a left sidebar with user options and a right sidebar with a "Content Browser" showing a hierarchical tree of categories.

Figura 6. Visualização de uma referência cadastrada, destacando os termos associados ao documento tanto na taxonomia individual como no modelo de referência.

2.3. Suporte à Web Semântica

A *Web Semântica* [Davies, Fensel e Van Harmelen 2003] é uma iniciativa ligada ao W3C³, um comitê gestor que visa a manutenção e o desenvolvimento constante da *World Wide Web*. Essa iniciativa tem o principal objetivo de criar e manter formatos direcionados a combinar e integrar dados advindos de diferentes sistemas de informação. Objetivando essa integração, são desenvolvidos padrões que possibilitam a compreensão semântica de informações por parte de programas. Desenvolvimentos na *Web Semântica* vêm permitindo que serviços *Web* sejam cada vez mais automatizados,

³ <http://www.w3.org/>

sejam eles serviços comerciais como a compra e venda de passagens e outros produtos, serviços de acesso à informação, etc.

Dentre os vários padrões difundidos pela *Web Semântica*, o RDF⁴ é um dos mais populares atualmente [Davies, Fensel e Van Harmelen 2003]. O sistema Biblioref faz uso de um microformato do RDF, chamado eRDF⁵ (embedded RDF), correspondendo a uma versão do RDF possível de ser codificada juntamente com o HTML de um website. Conseqüentemente, aplicações que compreendem esse microformato, também podem obter e processar as informações que estão disponíveis no sistema. Com isso, é possível que outros sistemas sejam capazes de efetuar buscas no sistema Biblioref sem precisar ter acesso a seu banco de dados ou a informações sobre sua organização interna. Nessa versão do sistema, todas as páginas que exibem conteúdo sobre referências bibliográficas possuem essa codificação.

Os navegadores *Web* tradicionais não processam essa informação, deixando a visualização das páginas como exibidas nas Figuras 5 a 6. Outros programas, porém, compatíveis com esse padrão, podem fazer uso da informação codificada em eRDF para oferecer serviços interessantes, como, por exemplo, aconselhar a alguém que entre em contato com um pesquisador que colecioner determinados tipos de documento, sugerir colaboradores baseado em similaridades de documentos, entre outros.

A Figura 7 traz um exemplo de código eRDF, ilustrando como a página exibida na Figura 6 é processada por um programa que interpreta eRDF (saída no formato XML de tal programa).

```
<rdf:RDF>
- <rdf:Description rdf:about="">
  <admin:generatorAgent rdf:resource="http://purl.org/NET/erdf/extract"/>
  </rdf:Description>
- <rdf:Description rdf:about="http://rmanola.890m.com/?q=content/wiki-content-templating">
  <rdfs:label>Wiki content templating</rdfs:label>
  <dc:title>Wiki content templating</dc:title>
  <dc:creator>Angelo Di Iorio</dc:creator>
  <dc:creator>Fabio Vitali</dc:creator>
  <dc:creator>Stefano Zacchiroli</dc:creator>
  <dc:identifier>/?q=content/wiki-content-templating</dc:identifier>
  <rdfs:isDefinedBy>Knowledge Management</rdfs:isDefinedBy>
  <rdfs:isDefinedBy>Wiki</rdfs:isDefinedBy>
  </rdf:Description>
- <rdf:Description rdf:about="http://rmanola.890m.com/?q=content/wiki-content-templating#">
  <rdf:type rdf:resource="http://xmlns.com/foaf/0.1/Document"/>
  </rdf:Description>
</rdf:RDF>
```

Figura 7. Exibição em eRDF da página da Figura 5

⁴ Resource Description Framework (RDF)

⁵ Mais detalhes sobre eRDF são encontrados em: <http://research.talis.com/2005/erdf/wiki/Main/RdfInHtml>

Como pode ser visto na Figura 7, o conteúdo do eRDF é exibido em formato XML. Assim, uma aplicação compatível pode saber quais são os autores do documento, qual é o endereço dessa informação e quais termos de classificação compõem sua organização.

3. Trabalhos Relacionados

A característica principal do Biblioref é permitir alto grau de autonomia ao usuário na organização, visualização e busca do conhecimento. Em outras palavras, o usuário é capaz de criar seu próprio esquema de classificação e o sistema provê um mecanismo permitindo associar essa informação a um modelo de referência. Existem outras iniciativas que fazem uso de taxonomias individuais para viabilizar a gestão do conhecimento. Dentre eles, os mais conhecidos são KARE [Guizzardi, Ludermir e Sona 2007] e KEEEx [Bonifacio *et al.* 2004].

KARE é um sistema de recomendação que usa taxonomias não apenas para classificar documentos PDF mas tudo o que é considerado um artefato de conhecimento. A partir dessa classificação, o sistema oferece respostas a perguntas em linguagem natural submetidas pelos usuários. Se, por exemplo, a resposta for encontrada no log de uma seção de bate-papo com um especialista, esse log é recuperado e posteriormente classificado na taxonomia do perguntador, tal como seus demais documentos. Como Biblioref, KARE permite ao usuário construir sua classificação própria, porém, ao invés de usar um modelo de referência, o sistema (KARE) oferece um algoritmo que mapeia as taxonomias dos usuários diretamente.

KEEEx também oferece um método automatizado para determinar a similaridade entre taxonomias. Tal como Biblioref, os usuários classificam seus documentos através de taxonomias denominadas contextos. Posteriormente, eles podem realizar buscas por outros documentos, baseando-se em palavras-chave ou em similaridades entre contextos. Tanto KARE quanto KEEEx implementam o modelo peer-to-peer enquanto Biblioref é baseado em uma arquitetura cliente-servidor.

A grande vantagem da abordagem do sistema Biblioref, ou seja, de usar uma taxonomia de referência para permitir relacionamentos entre taxonomias de usuários, está relacionada à eficiência do sistema. Soluções automáticas como as adotadas pelos sistemas previamente descritos tendem a ter alto custo em termos de processamento, além de nem sempre encontrarem termos correspondentes entre as taxonomias [Guizzardi, Ludermir e Sona 2007]. A estratégia semi-automática adotada é melhor tanto em relação ao processamento quanto na obtenção de resultados adequados. Por outro lado, algo que pode ser visto como desvantagem é que Biblioref exige maior envolvimento do usuário no processo de mapeamento, já que, muitas vezes, ele deve indicar qual o termo correspondente ao seu no modelo de referência. Entretanto, acredita-se aqui, que um bom equilíbrio foi atingido no sentido de garantir eficiência e regular o grau de esforço empregado pelo usuário no uso do sistema.

Caravela [Aumueller e Rahm 2007] também apresenta alguma similaridade com o trabalho aqui descrito. Ele também representa um sistema de gerenciamento de conteúdo com integração e categorização automática de informação. Nesse caso, o objetivo é ser utilizado como um repositório de conteúdo que classifica automaticamente o conteúdo submetido. Essa classificação também se baseia em

taxonomias que, porém, são centralizadas e não individuais. Usuários podem apenas postar conteúdo, que é compartilhado em um repositório central. Tanto Caravela quanto Biblioref são baseados no mesmo CMS, ou seja, o Drupal.

4. Discussão: Modelos Semânticos como Base para Sistemas Colaborativos de Gestão do Conhecimento

Taxonomias e ontologias são dois tipos de modelo semântico que podem ser aplicados para descrever determinado domínio, modelando-o como um conjunto de conceitos (termos) e relações. Ainda que tenham sido aplicados a sistemas de informação no passado, a Web Semântica tem aumentado o interesse em utilizar esse tipo de modelo conceitual para explicitar o conteúdo de artefatos de conhecimento [Davies *et al.* 2003]. Uma taxonomia pode ser vista como uma ontologia simplificada. Em outras palavras, ambos os tipos de modelo tratam conceitos (termos) da mesma maneira; entretanto, no que se referem às relações, taxonomias focalizam relações hierárquicas, enquanto ontologias podem exibir, além dessas, outros tipos mais complexos de relação (tais como relações todo-parte, relações de dependência, relações materiais etc.) [Guizzardi 2005].

Acredita-se que tanto taxonomias quanto ontologias sejam modelos bastante úteis para viabilizar a classificação de conteúdo em sistemas de suporte ao trabalho colaborativo e à gestão de conhecimento, bem como para favorecer a colaboração entre usuários desses sistemas. Isso se verifica na presente proposta, mas também em outros sistemas existentes [Guizzardi, Ludermir e Sona 2007; Guizzardi 2006; Bonifacio *et al.* 2004; Davies *et al.* 2003].

No sistema proposto neste trabalho, bem como em outros descritos na seção 3, taxonomias são adotadas com vistas a fornecer, como previamente discutido na seção 2, maior autonomia ao usuário com respeito ao esquema de classificação utilizado para organizar seus artefatos de conhecimento pessoais. Além disso, no sistema Biblioref, há também uma taxonomia de referência central, utilizada para relacionar as taxonomias do usuário. É justamente esse modelo que suscita a presente discussão. Apesar da versão corrente de Biblioref ser baseada em uma taxonomia de referência, é possível que esse modelo central seja substituído por (ou elaborado com base em) uma ontologia de domínio. Tal decisão se justifica porque, ao contrário das taxonomias individuais, que são modelos mais simplificados, o modelo de referência pode ser mais elaborado, já que é provido pelo sistema em si e, assim, deve ser fruto de trabalho especialista em determinada área de conhecimento. Tal modelo pode ser inclusive importado de repositórios e portais que disponibilizam ontologias de domínio existentes.

Algumas das vantagens que se podem intuir para a elaboração do modelo de referência a partir de ontologias de domínio: a) a obtenção de mais informação para a classificação de documentos (definição das relações discutidas na seção 2), a partir dos conceitos (termos) definidos por especialistas do domínio; b) a exploração dos tipos mais complexos de relacionamento de uma ontologia como meios de melhorar a classificação e busca de documentos; e c) o uso desse modelo para melhorar o mapeamento das taxonomias individuais, visando identificar similaridades que indiquem possibilidade de colaboração entre usuários. Assim, um dos trabalhos futuros relativos a esta iniciativa refere-se a elaborar e testar tal hipótese.

5. Considerações Finais

Este artigo apresenta Biblioref, um sistema colaborativo baseado em modelos semânticos para o compartilhamento de documentos. Tal iniciativa desenvolve idéias e estratégias de trabalho colaborativo e gestão de conhecimento abordadas em trabalhos anteriores, em particular em [Gomes 2006] e [Guizzardi 2006].

Como mencionado anteriormente, Biblioref é baseado no sistema Drupal. Este, por sua vez, é um software livre de API aberta, permitindo que o sistema aqui proposto fosse implementado de maneira mais eficiente, customizando o CMS existente para que o mesmo realizasse as funcionalidades desejadas. Como esse tipo de sistema apresenta certa complexidade, além de limitações na documentação disponível, tal customização não ocorreu sem alguma dificuldade inicial. Para viabilizá-la, foi preciso efetuar ampla participação em listas de discussão e bate-papo, bem como realizar leitura técnica aprofundada sobre o sistema.

Os primeiros testes com o sistema foram realizados com base no domínio de Trabalhos Colaborativos (CSCW), *i.e.* a taxonomia de referência testada se refere a essa área de conhecimento. Entretanto, o escopo do sistema pode ser estendido para outros domínios e aplicações, fundamentalmente dependendo da taxonomia de referência utilizada e do tipo de informação que melhor descreve os documentos compartilhados (*e.g.* resumo, no caso de artigos científicos; manchetes, no contexto de notícias *etc.*).

Os próximos passos desta pesquisa referem-se ao desenvolvimento de testes empíricos de *desempenho* e *usabilidade*, ambos muito importantes para verificar a eficácia da abordagem semi-automática de classificação de documentos. Deve-se, assim, avaliar, mais a fundo, o desempenho do sistema, principalmente, do ponto de vista da eficiência do algoritmo usado. Por outro lado, testes de usabilidade também são necessários para verificar como usuários reais avaliam os ganhos obtidos com o sistema desenvolvido. Em particular, a escolha entre oferecer maior eficiência e exigir maior esforço do usuário, comentado na seção 3, é bastante freqüente nesse tipo de sistema. Acredita-se que esses testes empíricos permitam um maior entendimento sobre como balancear estratégias automáticas e semi-automáticas de mapeamento de taxonomias.

Uma outra perspectiva de trabalho é adicionar novas funcionalidades ao Biblioref. Por exemplo, pretende-se desenvolver um método de busca por profundidade, que fornece informação contextualizada sobre o conteúdo classificado em níveis diferentes das taxonomias do usuário. O resultado traria documentos ranqueados tanto em termos desses níveis como em relação à similaridade quanto à consulta realizada pelo usuário.

Agradecimentos

Este trabalho foi parcialmente apoiado pelo Programa Institucional de Bolsas de Iniciação Científica (PIBIC), concedendo uma bolsa ao autor Renan Manola, e pela FAPES, através da bolsa DCR número 37274554/2007 concedida à autora Renata S.S. Guizzardi. O trabalho também foi parcialmente apoiado pelo projeto PPP (FAPES/MCT/CNPq/CT-INFRA nº 19/2006) processo nº 36316008/2007.

Referências

- Bonifacio, M., Bouquet, P., Mameli, G. e Nori, M. (2004) Peer-Mediated Distributed Knowledge Management. In: Agent-Mediated Knowledge Management, Editado por van Elst, L., Dignum, V., and Abecker, A. Springer- Verlag, Alemanha, p. 31–47.
- Davies, J., Fensel, D. e Van Harmelen, F. (2003) Towards The Semantic Web: Ontology-Driven Knowledge Management. Wiley, Inglaterra.
- Fischer, G. e Ostwald, J. (2001) Knowledge Management: Problems, Promises, Realities, and Challenges In: IEEE Intelligent Systems, v. 16, n. 1, Jan/Feb01.
- Gomes, R.L. (2006) LEICA: Un Environnement Faiblement Couplé pour l'Intégration d'Applications Collaboratives, Tese de Doutorado, Université Paul Sabatier, Toulouse, França.
- Guizzardi, G. (2005). Ontological Foundations for Structural Conceptual Models. Tese de doutorado, University of Twente, Holanda.
- Guizzardi, R.S.S., Ludermir, P.G. e Sona, D. (2007) A Recommender Agent to Support Knowledge Sharing in Virtual Enterprises. In: Agent and Web Service Technologies in Virtual Enterprises, Editado por Protogeris, N. Idea Group Publishing, EUA.
- Guizzardi, R.S.S. (2006) Agent-oriented Constructivist Knowledge Management. Tese de doutorado, University of Twente, Holanda.
- Liao, S. (2002) Knowledge Management Technologies and Applications: Literature Review from 1995 to 2002. In: Expert Systems with Applications, v. 25, p. 155-164.
- Merali, Y. e Davies, J. Knowledge Capture and Utilization in Virtual Communities. In: Proc. of the First International Conference on Knowledge Capture, Canada.
- Nonaka, I. e Takeuchi, H. (1995) The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation, Oxford University Press, EUA.
- Pumareja, D., Bondarouk, T. e Sikkil, K. (2003) Supporting Knowledge Sharing Isn't Easy – Lessons Learnt from a Case Study. In: Proc. of the Information Resource Management Association International Conference, USA.
- Aumueller, David; Rahm, Erhard "Caravela: Semantic Content Management with Automatic Information Integration and Categorization" E. Franconi, M. Kifer, and W. May (Eds.): ESWC 2007, LNCS 4519, pp. 729–738, 2007.