UNIVERSIDADE FEDERAL DE PERNAMBUCO
GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO
CENTRO DE INFORMÁTICA
2012.1

Uma ferramenta para Definição de Mapeamentos entre Vocabulários usados na publicação de Dados Linkados

PROPOSTA DE TRABALHO DE GRADUAÇÃO

Aluno: Gabriela Fernanda Silva de Vasconcelos (gfsv@cin.ufpe.br)
Orientador: Bernadette Farias Lóscio (bfl@cin.ufpe.br)

Recife, Abril de 2012

1. Contextualização

Na *Word Wide Web* contemporânea, *a*s informações, em sua grande maioria, são apresentadas em um formato de um textual através de uma interface amigável. Entretanto, tais conteúdos geralmente seguem regras apenas sintáticas, com objetivos de apresentação, não permitindo que se consiga facilmente extrair semântica dos mesmos. Dessa forma, os resultados das buscas realizadas na Web geralmente são imprecisos, uma vez que não é possível fazer uma distinção do valor semântico dos dados.

Para resolver tais problemas foi proposta a Web Semântica [Lee *et al* 2001], uma extensão da Web atual, onde as informações apresentam um significado bem definido e bem estruturado [Gonçalves 2012]. Com a Web Semântica as informações podem ser integradas, compreendidas e interpretadas pelos computadores, permitindo, dessa forma, refinar com maior precisão os resultados das pesquisas realizadas pelos usuários [Marcondes *et al* 2008.

Um dos fatores essenciais para que a Web Semântica torne-se uma realidade diz respeito ao uso de ontologias. Uma ontologia é algo que define conceitos, relações e restrições de forma compreensiva, explícita e formal de uma área de conhecimento especifico [Gruber 1995]. Em uma ontologia é necessário que se possa produzir o raciocínio inteligente por meio de mecanismos de inferência que darão suporte à manipulação do conhecimento explicitado.

Atualmente, é possível encontrar diversas ontologias definidas para um mesmo domínio, cada qual com suas particularidades de modelagem. Essa diversidade de ontologias torna-se um problema quando é necessário consultar dados armazenados em ontologias distintas e que dizem respeito a um mesmo domínio. Nesse caso, uma consulta definida sobre uma ontologia precisará ser reescrita em consultas a serem executadas nas outras ontologias que se deseja acessar [Campos *et al* 2009].

Para facilitar esse processo é comum o uso de mapeamentos entre ontologias, onde se tem como resultado um conjunto de expressões que relacionam os termos de uma ontologia aos termos de outra ontologia [Felissísimo 2005]. As informações contidas no mapeamento vão depender do tipo de vínculo semântico encontrado entre os elementos e do tipo de formalismo utilizado para representação dos mapeamentos [Campos *et al* 2009].

Recentemente, como parte do desenvolvimento da Web Semântica, surgiu o conceito de *Linked Data* (dados ligados) que pode ser definido como um conjunto de boas práticas para publicar e conectar conjuntos de dados estruturados na Web, com o intuito de criar uma "Web de

Dados" [Bizer *et al.* 2009]. A crescente adoção dos princípios de *Linked Data* e a evolução contínua da nuvem de dados linkados têm deixado ainda mais evidente a necessidade de definição de mapeamentos entre ontologias. Um dos princípios de *Linked Data* diz respeito ao uso de vocabulários existentes para a publicação de dados na Web, a fim de facilitar seu compartilhamento. Porém, a grande quantidade de vocabulários disponíveis (definidos por meio de ontologias) acaba levando à publicação de um mesmo conteúdo, em sites distintos, com termos distintos. Dessa forma, para ter acesso integrado à estes conteúdos é necessário que sejam definidos mapeamentos entre tais vocabulários. Neste trabalho, estamos interessados na definição de mapeamentos entre os principais vocabulários usados atualmente para a publicação de dados linkados na Web. A idéia da nossa proposta consiste em criar um repositório de mapeamentos entre vocabulários que possa ser compartilhado por diversas aplicações

2. Objetivos

O objetivo geral deste trabalho é a implementação de uma ferramenta para criação de mapeamento entres os vocabulários (definidos por ontologias) visando o reuso dos termos dessas ontologias não apenas na criação de outras ontologias, facilitando a comunicação das diferentes aplicações, pois existem termos, nos diversos vocabulários, com grafias diferentes, mas que semanticamente conceituam o mesmo recurso em um domínio, ou ainda, possuir grafias iguais mas com significados distintos, como também aprimorar os conhecimentos de quem estará se beneficiando da ferramenta porque terá a oportunidade de conhecer outras ontologias a partir do mapeamento com a sua ontologia de conhecimento sem precisar recorrer a uma pesquisa na internet, pois o mapeamento estará centrado em só lugar.

Com o uso da ferramenta proposta será possível além de criar o mapeamento entre ontologias, pesquisar os termos de cada ontologia, bem como a relação dos termos de uma ontologia com os termos de outra ontologia.

Como parte deste trabalho, será realizado um levantamento das principais soluções para armazenamento de dados para RDF (*Resource Description Framework*), uma vez que os mapeamentos entre as ontologias serão representados como triplas. Além disso, faremos uma análise das abordagens usadas atualmente para definição de mapeamentos entre vocabulários e sua aplicação no contexto de aplicações de integração de dados linkados.

3. Cronograma

O cronograma de atividades previsto para o desenvolvimento deste Trabalho de Graduação é apresentado na Tabela 1.

Atividade	Abril				Maio				Junho				Julho			
Levantamento do estado da arte e definição do escopo	X	X	X	X												
Estudo das abordagens definição de mapeamentos entre vocabulários				X	X											
Especificação da ferramenta						X	X	X								
Implementação do protótipo									X	X	X					
Validação e testes											X	X				
Melhorias no protótipo a partir dos resultados dos testes													X			
Elaboração do relatório final									X	X	X	X	X			
Defesa														X		

Tabela 1: Cronograma de atividades

5. Possíveis Avaliadores

Os possíveis avaliadores desse trabalho de graduação serão:

- Ana Carolina Salgado
- Bernadette Farias Lóscio

6. Referências

[Bizer *et al* 2009] Bizer C., Heath T., Berners-Lee T. (2009) *Linked data - the story so far*. Int. J. Semantic Web Inf. Syst., 5(3):1–22, 2009.

[Gruber 1995] Gruber T. (1995) *Toward principles for the design of ontologies used for knowledge sharing*. 1995. International Journal Human-Computer Studies Vol. 43, Issues 5-6, November 1995, p.907-928.

[Lee *et al* 2001] Lee, B. T., Hendler J., Lassilia O. (2001) *The semantic web*. Scientific American, 284(5):34–44, Mai 2001. http://dx.doi.org/10.1038/scientificamerican0501-34DOI: 10.1038/scientificamerican0501-34

[Gonçalves 2012] Gonçalves, Frederico Luiz de Freitas. Ontologias e a Web. Abril 2012. http://www.inf.ufsc.br/~gauthier/EGC6006/material/Aula%203/Ontologia_Web_semantica%20Freitas.pdf

[Marcondes *et al* 2008]. Marcondes, Carlos Henrique, Campos, Maria Luiza de Almeida (2008). Ontologia e Web Semântica: O espaço da Pesquisa em Ciência. Abril 2012. http://www.portalseer.ufba.br/index.php/revistaici/article/viewArticle/2669

[Freitas *et al* 2007]. Freitas, Juliano Baldex de, Lima, Vera Lúcia Strube (2007). Uma análise da avaliação dos resultados da medida de Similaridade Semântica (SiSe) entre ontologias em português. Abril 2012.
http://www.de9.ime.eb.br/~sousamaf/cd/pdf/arq0161.pdf

[Campos *et al* 2009]. Campos, Maria Luiza de Almeida, Campos Maria Luiz Machado (2009). Aspectos metodológicos no reuso de ontologias: um estudo a partir das anotações genômicas no domínio dos tripanosomatídeos. Abril 2012.

http://www.reciis.cict.fiocruz.br/index.php/reciis/

[Felissísimo 2005]. Felissísimo, Carolina Howard (2005). Interoperabilidade Semântica na Web: Uma Estratégia para o Alinhamento Taxonômico de Ontologias. Abril 2012. http://www.maxwell.lambda.ele.puc-rio.br/Busca_etds.php?strSecao=resultado&nrSeq=5864@1

7. Assinaturas

_____

Gabriela Fernanda Silva de Vasconcelos
Orientando(a)

_____

Bernadette Farias Lóscio
Orientador(a)

Recife, Abril de 2012