UNIVERSIDADE FEDERAL DE PERNAMBUCO

GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

CENTRO DE INFORMÁTICA

2011.1

**Mapeamento XML para Ontologias OWL**

PROPOSTA DE TRABALHO DE GRADUAÇÃO

**Aluno:** Philippe Pereira das Neves (ppn@cin.ufpe.br)

**Orientador:** Ana Carolina Salgado (acs@cin.ufpe.br)

Recife, 13 de Abril de 2011

Sumário

[1. Contexto 3](#_Toc290582401)

[2. Objetivos 4](#_Toc290582402)

[3. Cronograma 4](#_Toc290582403)

[4. Possíveis Avaliadores 4](#_Toc290582404)

[Referências 5](#_Toc290582405)

[Data e Assinaturas 6](#_Toc290582406)

# Contexto

Embora a WEB tenha sido projetada para possibilitar o fácil acesso, intercâmbio e a recuperação de informações, não há nenhuma estratégia abrangente e satisfatória para a indexação dos documentos nela contidos [1]. A recuperação das informações a partir dos motores de busca é baseada primariamente em palavras-chave contidas no texto dos documentos compartilhados, o que deixa a desejar na precisão das consultas realizadas. Além disso, as tecnologias e linguagens atualmente utilizadas nas páginas Web focalizam os aspectos de exibição e apresentação dos dados, de forma que a informação seja pobremente descrita e pouco passível de ser entendidas por máquinas e seres humanos [1]. Neste contexto, surge a proposta da Web Semântica.

O termo Web Semântica refere-se à visão da W3C (World Wide Web Consortium) que pretende melhorar a forma com que programas podem interagir com páginas Web, adicionando inteligência e contexto nos códigos XML. É um esforço colaborativo liderado pelo W3C com a participação de um grande número de pesquisadores e parceiros industriais [2].

O padrão XML é aceito como o padrão para troca de dados na Web. Mas, apesar de possibilitar aos autores a criação de suas próprias tags, em uma perspectiva computacional, há muito pouca diferença entre as tags <AUTHOR> e <CREATOR>. Para que as marcações semânticas criadas sejam utilizadas de forma não ambígua por comunidades maiores, são necessários alguns padrões de compartilhamento mais universais [1].

A Web Ontology Language OWL é uma linguagem de marcação semântica para publicação e compartilhamento de ontologias na World Wide Web [2]. OWL é desenvolvido como uma extensão do vocabulário de RDF (Resource Description Framework) e é derivado do DAML + OIL Web Ontology Language. A OWL é projetada para uso por aplicações que precisam processar o conteúdo da informação em vez de apenas apresentar informações, tornando mais fácil do que em XML, a interpretação por máquinas do conteúdo da Web, por fornecer vocabulário adicional com uma semântica formal [2].

# Objetivos

O objetivo deste trabalho é realizar o mapeamento do XML para OWL, de forma que ontologias OWL possam ser geradas automaticamente a partir dos elementos de esquema XML. Em particular, espera-se propor uma estratégia para identificar e mapear, elementos com referência a outros elementos no modelo XML, para o relacionamento correspondente em código OWL.

# Cronograma

A tabela 1 representa o cronograma das atividades a serem realizadas durante o trabalho proposto neste documento. A cada mês considerado, são associadas quatro sub-colunas, estas, por si, representam as semanas que compõem este mês. As linhas dizem respeito às atividades fundamentais à condução do projeto. Desta forma, cada célula pintada na tabela representa uma semana que será dedicada à realização da tarefa associado à linha em questão.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **ATIVIDADES** | **MARÇO** | **ABRIL** | **MAIO** | **JUNHO** |
| Pesquisa e definição do escopo |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| Desenvolvimento da solução |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| Realização do estudo de caso e testes |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| Preparação do relatório |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| Preparação da apresentação |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| Defesa |   |   |   |   |   |   |   |   |   |   |   |   |   |   |

**Tabela 1: Cronograma das Atividades**

# Possíveis Avaliadores

Os possíveis avaliadores para o resultado a ser obtido ao final de todas as etapas da proposta descrita neste documento são:

* Bernadete Loscio
* Patrícia Tedesco

# Referências

[1] Renato Rocha Souza, Lídia Alvarenga: A Web Semântica e suas contribuições para a ciência da informação, 2003. Disponível em: <http://www.scielo.br/pdf/ci/v33n1/v33n1a16.pdf>. Acesso em: Abril de 2011.

[2] World Wide Web Consortium (W3C): OWL Web Ontology Language Overview. Disponível em: <http://www.w3.org/TR/owl-features/>. Acesso em: Abril de 2011.

# Data e Assinaturas

Recife, 15 de agosto de 2011.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Ana Carolina Salgado

**Orientador**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Philippe Pereira das Neves

**Aluno**