

Proposta de uma Ontologia para DataWebhouse

Luis Henrique Bogo^{1,2}, Heleno Fulber^{1,2}, Raphael Winckler de Bettio^{1,2}, Gustavo Burckhardt^{1,2}, Alejandro Martins Rodriguez^{1,2}

¹ Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento - UFSC
Campus Universitário –Florianópolis – SC – Brasil

² Instituto Virtual de Estudos Avançados

Florianópolis – SC - Brasil

{bogo,heleno,gustavob,raphael,martins}@egc.ufsc.br

***Resumo:** A Web nasceu sob o signo das linguagens de marcação apenas para apresentação de seu conteúdo. Com o tempo, surgiram novas necessidades, entre elas a de se representar o conhecimento embutido nas páginas apresentadas. Esse conceito pode ser aplicado em um sistema de comércio eletrônico de uma maneira que vise mapear o comportamento de um usuário para a futura utilização destas informações. Este artigo tem como objetivo criar uma ontologia que permita a troca de informações entre empresas sobre a navegação de clientes em um sistema de comércio eletrônico via web.*

1. Introdução

Com a popularização da Internet, os sistemas de comércio eletrônico recebem milhares de acessos diários, tornando possível a coleta de informações valiosas que podem auxiliar na melhoria dos serviços oferecidos e a definição de estratégias competitivas. Mas, como as páginas Web atuais são destinadas à leitura, os computadores até podem interpretar as páginas no que se refere a rotinas de formatação, mas são incapazes de inferir o conteúdo semântico da página, pois os dados encontram-se escondidos dentro de arquivos HTML, em meio a um emaranhado de códigos específicos para a formatação da página. A Web semântica surgiu como o início de uma forma eficiente de se representarem dados na World Wide Web.

Este trabalho tem como objetivo a criação de uma ontologia que permita a troca de informações – da navegação de clientes em sistemas de comércio eletrônico – entre instituições.

2. DataWebhouse

O datawebhouse permite analisar todo o caminho realizado por um visitante em um site da

Web, permite mapear cada clique, conhecer melhor cada cliente via informação disponibilizada nos arquivos de log do servidor e cookies. Assim, torna-se muito mais fácil e real uma análise de cada evento realizado dentro do site pelo visitante (cliente), segmentando todas as informações para o uso adequado [Fulber, 2004].

3. Web-semântica

O termo Web Semântica foi cunhado por Tim Berners-Lee, visando incluir alguma inteligência e significado ao contexto de visão e navegacional da Web.

A fim de permitir que os agentes de software possam compreender o significado de uma informação sem que para isso tenham que realizar processamento de linguagem natural usa-se a marcação semântica, baseada em metadados [Antoniou, 2004].

Segundo [Silva, 2005], uma ontologia define termos para que um agente de software consiga extrair o máximo de informação possível de um documento. As ontologias permitem que possamos descrever formalmente um certo domínio de aplicação sob um ponto de vista específico, pois, normalmente, não estamos interessados em todos os pontos de vistas possíveis.

A linguagem que vem sendo desenvolvida para a descrição de ontologias é a WOL (*Web Ontology Language*), a qual baseia-se no RDF. As ontologias compõem-se como o principal componente da web semântica.

4. Estudo de Caso

Para demonstrar a possibilidade de união entre aplicações semânticas e DataWebhouse, foi-se desenvolvida uma ontologia que permite a troca de informações – da navegação de clientes em sistemas de comércio eletrônico – entre instituições. Esta ontologia – disponibilizada através de um arquivo OWL – permitirá que uma instituição mapeie o perfil (interesses) dos clientes de outra organização e dá subsídios para o entendimento de questões tais como:

- a) Qual a relação entre consultas e efetivação da vendas?
- b) Clientes com perfis semelhantes possuem interesses semelhantes?
- c) Em que faixa etária devem-se direcionar os esforços de marketing? Etc.

Para o desenvolvimento da ontologia, utilizou-se o sistema Protégé.

4.1 Ontologia proposta

A estrutura desenvolvida de classes da ontologia está apresentada na figura 01, retirada da ferramenta de modelagem Protégé.



Figura 01. Estrutura de classes

A primeira estrutura representa a dimensão Produto e mapeia o Departamento, a Seção e a Categoria dos mesmos. Para dimensão Cliente, foram mapeados Sexo e faixa etária. Já a classe Acesso representa os fatos propriamente ditos, mapeando qual Cliente, Consultou/Reservou/Comprou, qual Produto. Nas figuras 02 está apresentada uma instância de Acesso que utiliza as classes Cliente e Produto.

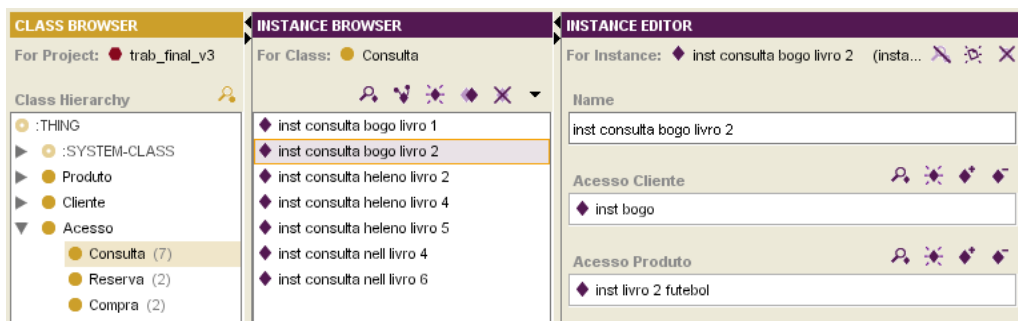


Figura 02: Instância de Acesso

Ao realizar pesquisas nos dados, principalmente quando forem analisados os “fatos”

de Acesso (Consulta – Reserva – Compra), o encadeamento de consultas se faz necessário.

5. Conclusões

Neste trabalho, o foco criado da ontologia foi mapear a navegação dos usuários nos produtos do sistema. Após a modelagem conceitual, o sistema Protégé facilitou o processo de desenvolvimento da ontologia. Criaram-se subsídios para o entendimento quanto ao comportamento dos clientes possibilitando respostas a questões como as abaixo listadas:

- a) Compras de Produtos do Departamento “Livros”;
- b) Consulta de Produtos pertencentes ao Departamento/Seção “Livros/Literatura Nacional”;
- c) Reserva de Produtos feita por Clientes de Sexo/Faixa Etária “Masculino/Homem de 21 a 40” referentes ao Departamento/Seção/Categoria “Livros/Espportes/Futebol”.

Finalizando, a utilização da web semântica se apresenta como uma possível abordagem ao serem utilizados os datawebhouses, pois disponibiliza não apenas o acesso a uma massa de dados, mas a sua semântica também.

6. Bibliografia

Antoniou, Grigoris e Harmelen, Frank V., A Semantic Web Primer, The MIT Press, 2004.

Fulber, Heleno. (2004) “Proposta de um Modelo de Sistemas de Informação Baseado em Dara Warehouse para Alavancar a Gerência do Acompanhamento de Alunos em Ambiente de E-Learning”, Dissertação de Mestrado, Programa de Pós-Graduação em Engenharia de Produção e Sistemas, UFSC.

Kimball, Ralph e Ross, Margy., The data warehouse toolkit, Editora Campus, Tradução de Ana Beatriz Tavares e Daniela Lacerda, 2002.

Silva, George Henrique. (2005) “Construção de Agentes Inteligentes para a Web Semântica”, <http://www.linux.ime.usp.br/~cef/mac499-04/monografias/ghsilva/>, junho.