

UNIVERSIDADE FEDERAL DE SANTA CATARINA
DEPARTAMENTO DE INFORMÁTICA E ESTATÍSTICA
SISTEMAS DE INFORMAÇÃO
PROJETOS I

**Uma arquitetura baseada na Web Semântica para Ferramentas Analíticas
(Resumo)**

Dhiogo Cardoso da Silva

Florianópolis, 15 de junho de 2006

1 Introdução

As atuais condições de competição em que vivem as empresas forçam-nas a buscar meios para valorizar a informação e conhecimento em seus negócios. Muitas vezes, os gestores organizacionais utilizam ferramentas analíticas, que atuam no nível informacional, para auxiliar suas decisões. Porém as ferramentas de análise (ou ferramentas OLAP) podem ser ineficazes visto que possuem algumas limitações como:

- a) Falta de extensão de capacidades exploratórias;
- b) Não oferecem suporte para definições de lógica de negócio;
- c) Não apresentam a informação de acordo com o perfil de usuário;

A fim de solucionar essas lacunas, desenvolveu-se uma tecnologia semântica para definir uma arquitetura integrada para essas ferramentas. Esse texto é um resumo do artigo *A Semantic Web based Architecture for Analytical Tools* e é parte da avaliação da disciplina *Projetos I* do curso de *Sistemas de Informação* da UFSC.

2 Inteligência de Negócio (Business Intelligence)

Business Intelligence (BI) é definido como um conjunto integrado de ferramentas que dão suporte a transformação dos dados em informação visando a tomada de decisão. A arquitetura típica de BI contém Data Warehouse (DW), ferramentas de ETL (Extraction, Transformation and Loading) e as ferramentas de análise. As ferramentas analíticas possuem algumas funcionalidades típicas, tais quais, slice (i.e. filtrar por assuntos); dice (i.e. selecionar um conjunto de dados); drill-up (i.e. navegar de um nível menos detalhado para um nível maior de detalhamento); drill-down (i.e. o inverso de drill-up); drill-across (i.e. analisar sobre outras perspectiva, ou mover de um cubo para outro qualquer).

3 Tecnologias da Web Semântica

O objetivo principal da web semântica é propiciar a descrição de recursos web para que sistemas computadorizados possam localizar e fazer análises dos conteúdos disponibilizados. Ontologias são um dos principais caminhos para se chegar a esse objetivo. Elas podem ser definidas como conceitualizações sobre um domínio particular. O conhecimento pode ser representado por ontologias através de classes, propriedades e seus relacionamentos. O conhecimento representado por ontologias podem ser explorados por motores de inferência para gerar conhecimento adicional. O sistemas desenvolvido utiliza a linguagem OCML como forma de representar o conhecimento e ainda, um conjunto de inferências ou regras para dar subsídio às ferramentas analíticas.

Outro componente chave capaz de realizar a interoperabilidade na Web é a tecnologia Web Service, que visa a integração de sistemas heterogêneos. A aplicação de ontologias para descrição de Web Service Semânticos e outras infra-estruturas dão facilidade para automatizar a descoberta e composição de serviços Web. Como Web Service Semântica desreve-se o IRS-III SWS Framework que oferece maneiras de transformar o código de programação em serviços web e ainda, invocá-los.

4 Um Arquitetura para Ferramentas Analíticas

Essa seção apresenta como os módulos da arquitetura estão dispostos. A figura 1 abaixo mostra toda a arquitetura.

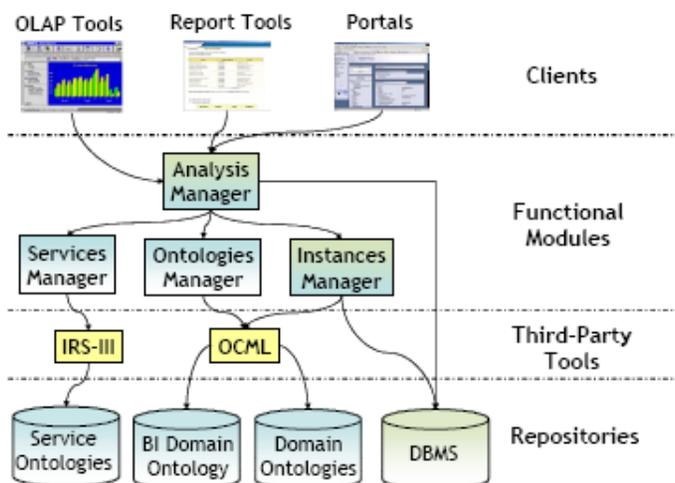


FIGURA 1 – Ilustração dos módulos da arquitetura
(fonte: artigo A Semantic Web based Architecture for Analytical Tools)

4.1 Ontologias de Domínio (Domain Ontology)

Essa ontologias oferece uma especificação formal das terminologias do domínio de negócio que é suportado pela arquitetura. Essa ontologia provê as terminologias de negócio que suportarão as descrições do web service semântica na Ontologias de Serviço (Service Ontology e contextualiza as fontes dos dados na Ontologia de Inteligência de Negócio (BI Ontology). Conseqüentemente, os usuários podem navegar pelos recursos (isto é, lógica, dados e serviços) usando conceitos do negócio em vez das descrições técnicas.

4.2 Ontologia de Inteligência de Negócio (Business Intelligence Ontology)

A ontologia de BI modela os conceitos usados para descrever como os dados estão organizados nas fontes de dados e para mapear esses dados

em conceitos descritos na Ontologia de Domínio. Essas definições são usadas para:

- Suportar inferências usando a Ontologia de Domínio para estender os resultados das consultas;
- Apresentar os resultados de acordo com o perfil do usuário;
- Guiar a atualização de instâncias na Ontologia de Domínio;

Através da linguagem OCML, utilizada no desenvolvimento da pesquisa, pode-se relacionar os conceitos do negócio com a fontes de dados, como um Data Warehouse, por exemplo. Com isso, consegue-se reescrever consultas à base de dados utilizando as extensões semânticas ou regras antes declaradas para explorar melhor os dados. Suponha que um usuário queira buscar produtos por categoria, sendo assim, a consulta que utiliza os conceitos da ontologia e da base de dados pode ser gerada através dos seguintes passos:

- Usando uma relação de transitividade descrita na Ontologia de Domínio, o motor de inferência encontra as superclasses desse produto e dos produtos derivados dessa superclasse;
- As chaves primárias dos produtos inferidos são incluídas na condição do comando 'select' ;
3. Outras restrições relacionadas a dimensão de produto são removidas da consulta;
4. Uma nova consulta é submetida à base de dados

As relações de transitividade são usadas pelas ferramentas OLAP para

suportar as operações de drill-down e drill-up sobre os resultado das análises.

Relações mais sofisticadas baseadas em expressões lógicas podem suportar maior poder de inferência sobre os dados. Como por exemplo, pode-se fazer as seguintes questões, Qual foi o rendimento de meu competidor no último quarto de ano? Sendo competidor uma relação especificada como “companhias que vendem o mesmo tipo de produto na mesma região”. Essa relação é expressa em OCML da seguinte forma como mostra listagem 1 a seguir.

Listagem 1. A relação competidor
(def-relation company-competitors (?c1 ?c2)
:sufficient (and (region ?r)
(product ?p)
(is_located_region ?c1 ?r)
(sells_product ?c1 ?p)
(is_located_region ?c2 ?r)
(sells_product ?c2 ?p)
(not (= ?c1 ?c2))))

Esse processo da listagem 1 e a atualização de instâncias na Ontologia de Domínio são guiados por uma série de classes que mapeiam os conceitos de domínio e banco de dados. Dessa forma os conceitos de DW como Fatos e Dimensões dentro outros são mapeados para classes respectivas.

A ontologia de BI também compreende conceitos usados para suportar a representação da estrutura dos dados e resultados das consultas de acordo com o perfil do usuário.

4.3 Gerenciador de Instâncias (Instance Manager)

O Gerenciador de instâncias suporta a replicação de dados de dimensões para a Ontologia de Domínio. Ele é guiado pelos mapeamentos descritos na Ontologia de BI. Nesse módulo há um conjunto de classes responsáveis pela conexão e extração dos dados a partir de uma base de dados.

4.4 Gerenciador de Ontologias (Ontology Manager)

O gerenciador de ontologia é o módulo que proporciona uma série de métodos necessários para consultar e realizar inferências sobre um repositório de ontologias. Seu objetivo tornar transparente para o usuário a complexidade de linguagens de consultas para ontologias, integrando as ferramentas analíticas, manipuladores de consultas e bases de ontologias.

4.5 Ontologia de Serviço (Service Ontology)

Modela os conceitos utilizados para descrever Web Services Semânticos (SWS). Atualmente, essa ontologia corresponde a um conjunto de ontologias usadas no framework IRS-III. Utiliza os conceitos da Ontologia de Domínio para descrever as entradas e saídas de framework.

4.6 Gerenciador de Serviço (Service Manager)

O objetivo desse módulo é fornecer a capacidade de reusar códigos a fim de propor melhorias na funcionalidade dos instrumentos de análise. Ele suporta a recomendação do acordo do SWS para combinar os conceitos envolvidos nas análises e nas descrições semânticas desses Web Services.

4.7 Gerenciador de Análises (Analysis Manager)

O Gerenciador de Análises é o módulo que fornece acesso a todos os

componentes da arquitetura. Serve como caminho para acessar o Gerenciador de Ontologias, Gerenciador e de Serviços e o Gerenciador de Instâncias.

5 OntoDSS – Uma ferramenta analítica para Suporte a Decisão

OntoDSS é uma ferramenta de front-end desenvolvida para validar a arquitetura de BI. Ela possui dois módulos principais: a) Definição de Análises (Analysis Definition) e b) Agente de Exploração (Exploration Recommender). O módulo de Definição de Análises dá suporte a usuários na definição de conteúdo e formato de suas análises. O vocabulário usado para representar as definições do usuário é inferido através de relações entre a base de dados e os perfis dos usuários na Ontologias de BI. Já o outro módulo , Exploration Recommender, oferece alternativas para exploração de resultados de análises para usuário. Além das funcionalidades tradicionais de OLAP, outras que confiam na semântica definida nos repositórios do ontologias.

6 Conclusão

As tecnologias que envolvem a Web Semântica juntamente com os conceitos de Ontologias têm sido aplicadas em diferentes formas para apoiar e melhorar algumas deficiências relacionadas aos sistemas de informações, mas não no contexto de ferramentas analíticas. Essas ferramentas atuais não oferecem formas escaláveis para agregar novas funcionalidades e ainda exibirem as informações de acordo com o perfil ou configurações do usuário. Através da arquitetura proposta é possível solucionar essas lacunas incorporando funcionalidades semânticas às ferramentas OLAP.