

1 Introdução

1.1 Motivação

Com a evolução da tecnologia, o homem vem contando com o auxílio cada vez mais poderoso de máquinas para realizar as mais diferentes tarefas. Acompanhando esta evolução, o volume de dados disponíveis vem crescendo exponencialmente, principalmente alavancado pela Internet. Sistemas de Gerência de Bancos de Dados (SGBDs) tornam-se então importantes para manipular e gerenciar tal volume de dados.

Os dados armazenados nos SGBDs são provenientes de empresas privadas e agências governamentais, entre outras fontes. Frequentemente, diversas destas fontes se referem a um mesmo domínio de aplicação como, por exemplo, geoprocessamento. São diversas as empresas privadas e agências governamentais que coletam e manipulam dados geográficos para análise das informações relevantes que estes dados podem fornecer. Porém, devido às diferenças nos requisitos das aplicações e na falta de padronização da modelagem dos bancos de dados, o resultado é a adoção de modelos de dados distintos para cada fonte de informação. Esta heterogeneidade dificulta sobremaneira a interoperabilidade entre os sistemas.

De fato, o compartilhamento de dados é um fator crítico para o sucesso das organizações devido à emergente necessidade de comunicação com seus parceiros de negócios. Entretanto, a heterogeneidade das fontes é uma das barreiras para que sistemas de informação possam interoperar.

A abordagem tradicional proposta pela comunidade de banco de dados para o problema da interoperabilidade consiste em identificar e tratar os relacionamentos inter-esquemas (mapeamento de esquemas). Uma forma de atacar o problema utilizando esta abordagem seria fazendo o mapeamento entre pares de esquemas conceituais. Porém, dependendo da amplitude que se deseja alcançar com a solução, isto seria inviável devido ao grande número de mapeamentos diferentes necessários. Para contornar esta questão, a abordagem mais comum consiste na criação de uma visão ou esquema

conceitual global para unificar os elementos correspondentes dos diversos esquemas conceituais. Assim, os esquemas conceituais originais devem ser mapeados para o esquema conceitual global.

Porém, a criação de um esquema conceitual global impõe uma limitação muitas vezes inaceitável num ambiente tão dinâmico como a Web, onde o número de novas fontes que surgem é crescente. Neste contexto, o ideal é garantir a independência das fontes, viabilizando que novas fontes integrem-se à comunidade e outras desliguem-se desta sem prejuízos para nenhuma das partes, nem a necessidade de adaptações, como na arquitetura de mediadores (vide seção 2.3.3).

1.2 Solução proposta

Esta dissertação introduz o conceito de Catálogo de Objetos baseado em Ontologia (*Ontology-based Object Catalog* - OnOC) e propõe uma arquitetura para OnOCs como uma estratégia para endereçar o problema de interoperabilidade entre fontes de dados heterogêneas. Como prova de conceitos desenvolveu-se um protótipo de um OnOC utilizando ontologias de exemplo construídas seguindo a metodologia proposta.

Para a construção de OnOCs, optou-se por uma solução híbrida, utilizando conceitos oriundos de abordagens tradicionais e inovadoras, tais como: catálogos de metadados, servidores de ontologias, mediadores e federações de bancos de dados (vide Capítulo 2).

Um OnOC oferece recursos para definir um vocabulário comum que facilite a comunicação entre os membros da federação. Este vocabulário é chamado de *ontologia de referência* e é o componente central do ambiente proposto. A interoperabilidade é viabilizada através dos mapeamentos definidos entre a ontologia de referência e ontologias descrevendo as fontes de dados, à semelhança da abordagem tradicional de banco de dados.

1.3 Organização da dissertação

O Capítulo 2 resume as barreiras para interoperabilidade entre sistemas de informação e as abordagens propostas para integração, incluindo, entre outras: as federações de bancos de dados, a arquitetura de mediadores, o uso de ontologias e repositórios de metadados.

O Capítulo 3 introduz a metodologia utilizada para viabilizar a interoperabilidade entre fontes de objetos através de um catálogo de objetos baseado em ontologias (OnOC). Este capítulo apresenta as situações de uso e os requisitos para OnOCs.

O Capítulo 4 define a arquitetura proposta para OnOCs, descrevendo os módulos necessários e os serviços essenciais para cobrir os requisitos levantados no Capítulo 3. A definição de uma arquitetura padrão para OnOCs visa facilitar implementações de OnOCs.

O Capítulo 5 descreve um detalhamento da arquitetura, mostrando os diagramas de casos de uso, pacotes, classes e seqüência. Além disso, são mostradas as operações básicas implementadas pela camada de interface provida pela arquitetura proposta, estas operações cobrem os serviços essenciais descritos no Capítulo 4.

O Capítulo 6 descreve o protótipo desenvolvido como prova de conceito da metodologia e da arquitetura introduzidas nesta dissertação e serve como referencial para criação de OnOCs. As ontologias de exemplo utilizadas nos testes realizados foram construídas com base na metodologia descrita no Capítulo 3.

Finalmente, o Capítulo 7 contém as conclusões e sugestões para trabalhos futuros.