

## 5 Conclusões e trabalhos futuros

Esta dissertação abordou, inicialmente, o problema de localização de dados geográficos. Mostrou que para tornar útil a grande quantidade de dados disponíveis atualmente é preciso armazenar sua descrição em catálogos de metadados.

Mostrou que o processo de obtenção da descrição de um dado geográfico poderia consumir muito tempo e, portanto, deveria ser automatizado. Para isso, apresentou uma estratégia de descrição automática de dados geográficos a partir de um dicionário geográfico, onde a idéia básica foram as coordenadas geográficas associadas aos dados que poderiam ser correlacionadas com entradas dos dicionários. Apontou que tais dicionários deveriam fornecer interface de consulta, via programa, capaz de identificar relacionamentos topológicos do tipo “contido em”, “atravessa”, “tangencia”, entre as coordenadas dos dados geográficos e suas entradas.

A abordagem descrita nesse trabalho é semelhante à utilizada em *The Role of Spatial Relations in Automating the Semantic Annotation of Geodata*, onde Klien e Lutz [2] utilizam um processo semi-automático para geração de descrições de dados geográficos. Eles propuseram operações GIS sobre dados geográficos novos e dados de referência, e ontologias geo-espaciais para gerar descrições dos dados. Analogamente, propusemos um método para identificação de objetos geográficos relacionados a dados geográficos utilizando operações GIS oferecidas por dicionários geográficos. Entretanto, no método de Klien e Lutz somente parte da descrição é construída automaticamente. Argumentamos que, para compartilhar grande volume de dados, o processo de descrição dos dados deve ser totalmente automático.

Em *The Role of Gazetteers in Geographic Knowledge Discovery on the Web*, Ligiane Souza et al. [3] também apresentam mecanismo de catalogação automática de dados baseado em dicionários geográficos, entretanto, como sugerem aplicar tal procedimento para dados de natureza geral como por

exemplo, documentos HTML, não é possível valer-se das características geoespaciais como forma de identificação. Seu método baseia-se na identificação de nomes geográficos no conteúdo dos dados a serem catalogados.

Portanto, dada a importância da automação do processo de catalogação de dados geográficos no compartilhamento e disseminação desses dados, o trabalho buscou apresentar uma arquitetura de software capaz de identificar dados geográficos, criar descrições a seu respeito e catalogá-los automaticamente. Como prova de conceito apresentou um projeto detalhado para a implementação e testes de uma aplicação com esse propósito.

Por tratar-se de problema semelhante, a arquitetura de software desenvolvida para catalogação de dados geográficos foi generalizada de forma a produzir um *framework* para catalogação automática de dados de quaisquer naturezas.

Alguns pontos do problema de geração automática de metadados, apresentados na dissertação, não foram, porém, considerados no protótipo e podem ser considerados em trabalhos futuros. A escala e a classificação do dado geográfico deveriam ser utilizadas para selecionar entradas do dicionário compatíveis com o dado geográfico. Como os dados utilizados na prova de conceito eram fictícios e o dicionário ADL Gazetteer não dispõe de informações sobre escala essas validações não foram implementadas.

Outro problema surge quando vários dicionários são consultados para obter-se metadados. Entradas análogas em dicionários diferentes devem ser identificadas para não produzirem duplicidade de informação.

Finalmente, formatos realistas de dados geográficos como, por exemplo, GeoTIFF e ESRI Shapefile não foram considerados.