Spatial Data Infrastructures: reuse beyond the metadata

Jugurta Lisboa Filho and Odilon Corrêa da Silva

Departamento de Informática - Universidade Federal de Viçosa (UFV)
CEP 36570.000 – Viçosa – MG – Brasil

jugurta@ufv.br, odilon.correa@gmail.com

Abstract. *Environmental concerns have become increasingly important to all sectors of society, but especially to sectors working directly with environmental preservation. Environmental area characterization requires the knowledge on natural resources and the understanding of their interactions and correlations. In this article we present our views on the need to add semantic knowledge to improve the recovery process and increase sharing of data and information on Spatial Data Infrastructures.*

Introduction

The rapid evolution of the World Wide Web prompted the development of a whole new class of information systems with different architecture from their predecessors. This movement is spreading to spatial information systems, allowing users to search, view, combine, query, and analyze spatial data over the Internet (Athanasis et al., 2008).

Dissemination and sharing of spatial data sets from different backgrounds can undergo a huge expansion if supported by a computing environment in which data is shared freely and in an integrated way. Davis et al. (2005) argue that this is not the current reality, in which access to information has been one of the greatest difficulties of every researcher, user, developer or specialist who deals with computational representations of space. One solution to minimize this problem is to use Spatial Data Infrastructures (SDI), enabling cooperation and sharing of data (Maguire and Longley, 2005).

The open and distributed character of SDI is nevertheless an aggravating factor in recovery and sharing of geospatial data and information (Athanasis et al., 2008). In general, search for services in a SDI is mainly based on keywords, spatial coordinates, thematic or temporal classification. With the search result the user can access the data to view or download it, analyze metadata or start a new search. This approach has a number of difficulties, especially for inexperienced users, who may not know which keywords to use, how to fill out properly query forms, or define the number of criteria to use (Hochmair, 2005).

Such limitations in knowledge recovery can become more difficult for complex searches, for instance, in environmental data searches. The knowledge about the produced data must therefore be available to researchers in order to allow them to analyze the characteristics and interactions of these elements.

Position

Taken these questions into account, we sought a new methodology to recover and share geospatial knowledge, which, together with the issues arising from the interaction between SDI components, is a central theme of the Geospatial Semantic Web (Egenhofer, 2002). Thus, besides the actual data, users need semantic information on these data, which becomes possible by transforming SDI into a broader infrastructure. We will contribute to this debate by proposing an enhanced SDI with analysis patterns and domain ontologies. This new mechanism is being called Geospatial Knowledge Infrastructure (GKI) (Abdelmoty, 2007). Our position is based on results obtained from recent work on knowledge reuse through the integration of

analysis patterns and domain ontologies (Silva, 2008a, Silva, 2008b, Silva, 2008c and Tavares, 2008). Next, we will discuss two points that are relevant to the debate.

(1) Knowledge Reuse

Fowler (1997) defines a pattern as: “An idea that has been useful in one practical context and will probably be useful in others”. Reuse of analysis patterns is proposed in other application domains. According to Lisboa et al. (2002), geographical application domains can benefit from a pattern-based approach, since several sets of geographical phenomena are recurrent in many applications.

Ontologies are written as a set of definitions of formal vocabulary, establishing properties to enable knowledge reuse, preventing rework or rediscovery of equivalent terminology (Guarino, 1998).

Numerous approaches with the use of ontologies have been proposed within the context of SDI. Athanasis et. al (2008) argue that ontologies can be used within the context of SDI to facilitate discovery and sharing of knowledge.

Furthermore, the use of ontologies is one of the most promising mechanisms to achieve semantic interoperability among geographical data sources within a SDI (Lutz, 2008). However, the development of integration mechanisms based on analysis patterns and ontologies with metadata would make a more comprehensive SDI, allowing not only data recovery, but also helping designers from several organizations and sectors of society to reuse patterns for conceptual modeling of geographic databases (Silva et al., 2008a).

(2) Sharing Knowledge

Technological and semantic compatibility are interesting for all participants of a geographic data sharing initiative (Kok, 2005). The Open Geospatial Consortium (OGC) coordinated one of the most recent initiatives to promote interoperability, proposing patterns for the use of services in geographic information sharing (Nogueras-Isso et al., 2005).

Silva et al. (2008b) describe a Geospatial Knowledge Infrastructure (GKI) architecture for sharing data and metadata on the Antarctic ecosystem. The approach presented in this project emphasizes knowledge reuse stored in different forms, such as: (1) data, (2) theories, (3) methodologies, (4) software and (5) models, which may be related to each other. The proposed architecture was specified in a generic way and can be extended and used in other SDI such as the environmental context of the Amazon.

Conclusion

In this work, we presented some important issues so that the knowledge produced by a community may be available to users, researchers and others, allowing them to analyze these elements' characteristics and interactions. This will not be an easy task, because in addition to gathering competence from different disciplines, one has to produce mechanisms that are influenced by ontologies in the process of discovery and sharing of information in an SDI.

Acknowledgements

This project is partially financed by FAPEMIG and CNPq/MCT/CT-INFO.

References

Abdelmoty, Alia (2007) Geospatial Knowledge Infrastructures Workshop. Welsh e-Science Centre (WeSC): Cardiff University. Available in <http://www.nesc.ac.uk/esi/events/832/>

Athanasis, N., Kalabokidis, K., Vaitis, M., Soulakellis, N. (2008) Towards a semantics-based approach in the development of geographic portals. Computers and Geosciences. In press (doi: 10.1016/j.cageo.2008.01.014)

Davis Junior, C. A., Souza, L. A. and Borges, K. A. V. (2005b) Disseminação de Dados Geográficos na Internet. In: Casanova, M. A. et al. (Org.). Bancos de Dados Geográficos. Curitiba: EspaçoGeo, p. 353-378.

Egenhofer, M. (2002). Toward the Semantic Geospatial Web. Tenth ACM international symposium on Advances in geographic information systems. McLean, Virginia, ACM Press.

Fowler, M. (1997) Analysis Patterns: Reusable Object Models. Menlo Park, CA: Addison Wesley Longman.

Guarino, Nicola. Formal Ontology and Information Systems. In: Proceedings of the First Int. Conference on Formal Ontology in Information Systems, Trento, Italy, Junho 1998.

Hochmair, H., 2005. Ontology matching for spatial data retrieval from Internet portals. In: Proceedings of Geospatial Semantics. Lecture Notes in Computer Science 3799, Mexico City, Mexico, pp. 166–182.

Kok, B., Loenen, B. (2005) How to acess the success of national spatial data infrastructures? Computers, Environment and Urban Systems, N.29, p.699-717

Lisboa Filho, J.; Iochpe, C.; Borges, K. A. (2002) Analysis patterns for GIS data schema reuse on urban management applications. CLEI Electronic Journal, v.5, n.2, p.1-15.

Lutz, M., Sprado, J., Klien, E. Schubert, C., Christ, I. (2008) Overcoming semantic heterogeneity in spatial data infrastructures. Computers and Geosciences. In press. (doi: 10.1016/j.cageo.2007.09.017)

Maguire, D. J.and Longley, P. A. (2005) The emergence of geoportals and their role in spatial data infrastructures. Computers, Environment and Urban Systems, N. 29. p 3-14.

Nogueras-Iso, J., Zarazaga-Soria, R., Béjer, R., Álvarez, P. J. and Muro-Medrano, P. R. (2005) OGC catalog services: a key element for the development of spatial data infrastructures. Computers and Geosciences, N. 31. p. 199-209.

Silva, Evaldo O. et al. Improving analysis patterns in the geographic domain using ontological meta-properties (2008a) Int. Conf. on Enterprise Information Systems (ICEIS), Barcelona, Espanha. p.256-261.

Silva, Odilon et al. Uma infra-estrutura de conhecimento geoespacial para compartilhamento de dados e metadados sobre o ecossistema antártico (2008b) In Congresso de tecnologias para gestão de dados e metadados do Cone Sul, Curitiba.

Silva, E. O.; Lisboa Filho, J.; Oliveira, A. P. Revisando padrões de análise para bancos de dados geográficos com meta-propriedades ontológicas (2008c) In Workshop on Ontologies and Metamodels in Software and Data Engineering (WOMSDE-SBES). Campinas. P 97-106.

Tavares, D. B. et al. Validação de Diagrama de Classes por meio da Técnica OntoCon (2008) Congreso Latinoamericano de Informática (CLEI), Santa Fé, Argentina. p.330-339.