Development of a semantic network for environmental geodata

Stefan Göbel¹ and Uwe Jasnoch¹

Abstract

The more information and data² are produced in the actual information society, the more important become mechanisms and systems which organize data and include information where to find which data. Most popular peculiarities of such information systems are digital archives, metadata³ information systems (MIS) and catalogue systems (CS). These systems are used as preparatory step/tool for any other domain-specific applications such as site planning projects.

Recent initiatives –both on national and international basis- in the area of spatial libraries and environmental MIS/CS provide access to a wealth of distributed data, e.g. necessary data in the context of site planning. On the other hand they offer only basic levels of interactivity and assistance to the user. This includes all steps of the information retrieval process: Query formulation, query modification, comparison of (metadata) result sets and search result presentation. Referring to this lack of usability, there are special action lines and topics in the actual EU IST programme, e.g. key action III.4.2 information visualization which aims to facilitate information access, data handling and navigation in (often unfamiliar) information spaces.

Based on that situation this paper describes the establishment of a semantic network for geodata as part of the project “Development of the InGeoForum Information Center” which represents a theme-overlapping metadata information system for geodata on national basis and enables public access to environmental information and networks. Here, different information visualization techniques are used to facilitate the information retrieval process and access to heterogeneous geodata. Thus both municipal authorities and planners as operative groups in the site planning process are supported. Further on these techniques enable the participation and involvement of interested citizens.

Chapter 1 exemplarily proposes the site planning process as one environmental application area and indicates the necessity of metadata information systems, semantic networks and information visualization techniques as instruments to locate appropriate

¹ Fraunhofer Institute for Computer Graphics, GIS department, Rundeturmstr. 6, 64283 Darmstadt, Germany.
² Between 80 and 85% of all data is ‘spatial’, e.g. coordinates, geographic entities or addresses.
³ Metadata or „data about data“ describe the contents, quality, origin and other characteristics of data.
geodata and to interact with it using graphic-interactive concepts. Chapter 2 presents a state-of-the-art analysis of existing approaches concerning information visualization and metadata information systems. Chapter 3 introduces new methods and concepts for the establishment of a semantic network for geodata and describes the integration of these concepts into the InGeo Information Center. Finally some useful hints show possibilities for further research as well as further integration and transmission possibilities into other environmental/geodata application domains.

1. **Environmental data in the site planning process**

Figure 1 shows the complex site planning process as one fundamental application area of environmental (geo)data: A lot of different parties are involved and have their individual tasks and various geodata is needed in the different steps of the site planning process. In order to carry out site planning according to law, municipal authorities need specific data together with expert knowledge. Planning engineers analyze the completeness, quality, accuracy and other characteristics of the environmental data (sometimes appropriate data is missing and has to be acquired by surveyors before) and prepare reports which are used as basis for decision-making by municipal authorities. Further on the results (plans, maps, etc.) have to be published due to environmental information and documentation laws.

Hereby, metadata information systems and various information visualization techniques could facilitate all phases of the multi-step site planning process.

![Figure 1](image_url)

Multi-step site planning process and necessary geodata
2. Information visualization in metadata information systems

Today, depending on region and country, geographic data are available with heterogeneous quality, format and accessibility. Geographic data are only to a limited extent documented through metadata and accessible through catalogue or metadata information systems, which have become very popular in recent years. Besides, tools for domain-specific analysis, planning and visualization for the study of traffic or environmental problems exist.

However, existing tools and information systems only offer basic functionality and there is still a lack of usability. Users are not provided by any visual feedback during the information retrieval process. Thus, it is not possible to compare/analyze results or to get an impression how to redefine a query or how to interpret results for decision-making.

In order to overcome these obstacles, in the last years different approaches settled in the field of information retrieval (databases, query formulation, visual query language), information visualization (visualization of networks and hierarchies, visualization of queries and documents, visual data mining) or mathematics (ontologies, conceptual analysis) could be taken into account.

Examples of such information systems are: VisDB (Database exploration using multidimensional visualization), Ontobroker (ontology-based search engine), TOSCANA (analysis instrument for classification systems to locate synonyms in keyword searches), InfoCrystal (A visual tool for information retrieval) or Quantum (A query and analysis Tool for use with metadata). Further examples are scientific projects such as GeoLens, VISAGE (an information-centric workspace), IA (Intelligent Archive), EOSDIS (metadata information system for satellite images), GRADI (A Graphical Database Interface for a Multimedia DBMS) and the Alexandria Digital Library. With regard to the establishment of MIS for geodata and environmental information systems on European basis, the EU funded projects such as GELOS, Web-CDS, ESMI, VIRGO, CITUI or EEIS have to be mentioned. All of these systems aim to facilitate (public) access to environmental data and information systems.

3. Semantic network for geodata

Whereas fundamental MIS and information visualization techniques primarily support “casual users” in order to locate appropriate data and to interpret search results or plans, other mechanisms and profound knowledge are necessary to facilitate specific tasks of planning engineers involved in the site planning process. Here, both visual data mining techniques as well as extendable context repositories for different application domains are developed in the InGeo IC\textsuperscript{4} project.

\textsuperscript{4}InGeo IC: Development of the InGeoForum Information Center, see http://www.ingeovic.de
Figure 2 illustrates the development of the semantic network for geodata, which integrates the different domains of geodata disciplines and stores semantic relations among geodata archives, metadata (formats) and theme-specific terminology such as domain values, valids or thesauri as well as tasks respectively steps of geodata applications.

Hereby, the ISO theme codes\(^5\) are settled in the centre of the semantic network. All geodata archives, metadata formats or keywords are related to these thematic terms and are integrated into the network. Apart from the thematic section of the network containing content-related semantic relations between terms and theme codes, the network includes a geo component, too. Here geographic-topologic relations are stored to indicate spatial relations between geographic names and entities. Figure three shows a snapshot of the semantic-related part of the network for geodata.

---

\(^5\) ISO theme codes: classification of geodata disciplines, field in the ISO/TC 211 metadata standard CD 19115.3 geographic information – metadata.
The establishment of the semantic network enables users to access the various geodata archives and geodata disciplines in different ways containing both traditional form-based search methods for keyword searches or geographic names and new innovative approaches such as the thematic and application-driven access.

Especially the thematic access shown in figure 4 supports casual users to locate appropriate geodata in the (sometimes) unfamiliar field of heterogeneous geodata application areas.

5. Summary

This paper describes several aspects concerning metadata and metadata information systems as search instrument to locate appropriate geodata for complex geodata applications such as the multi-step site planning process. In order to enhance usability of these information systems both visualization techniques and semantic access variants could be integrated. The description of a the semantic network of geodata offers a good possibility to provide natural search methods containing an abstract theme-related access based on the ISO theme codes.
Concerning the establishment and extension of the semantic network, further research could concentrate on the usage of knowledge discovery and knowledge management techniques with regard to the semi-automatic integration of new terms and the generation of new semantic relationships.

Figure 4
Thematic access to geodata

References

