GeoShare – Building a Transnational Geodata Infrastructure for the North Sea Region

Thomas Vögele, Rainer Spittel, Ubbo Visser, Sebastian Hübner

Center for Computing Technologies (TZI)
University of Bremen

{vogele | rasp | visser | huebner}@tzi.de

SUMMARY

Funded by the EU Interreg IIIB North-Sea Region Programme, a transnational geodata infrastructure will be implemented within the framework of the GeoShare project. With 5 partners from 4 European countries (Norway, The Netherlands, United Kingdom, and Germany), the project is focused in the thematic areas of natural/cultural heritage and tourism, labour and education, and spatial planning. The overall goal of the project is to enhance the access to and exchange of geodata pertaining to these application areas. For this purpose, a services-based geodata infrastructure will be implemented, providing a common platform for the implementation and integration of regional applications. The development of this platform will be based on concepts and specifications issued by the international OpenGIS Consortium. The platform will be built using standard OGC-compliant open source and commercial software where possible as well as innovative tools developed by the project partners, like for example the intelligent catalog service that will be the core of the GeoShare Network.

INDRODUCTION

European Government policy agendas of modernization, partnership working, integrated service provision and democratic accountability have to be supported by communication and information systems with the ability to share information across organizational boundaries. Much progress is being made in individual service areas and within particular public service provider organizations. However, in many cases these organizations are not fully utilizing the potential of communication and information technologies. In particular this is true when it comes to utilizing the growing stores of geospatial data held within the public sector. Experts estimate that there is a growing market potential for the use of these data, while at the same time, adequate information about their content and availability is still missing (Greve and Plümmer 2002).

In order to utilize this information for a number of purposes ranging from educated decision making to inform the public, infrastructures for the effec-

tive distribution and access to geo-referenced data have to be created on a regional, national, and international level. In Germany, examples for such infrastructures are the GDI-NRW (Bernard 2002) and the evolving national geodata infrastructure (IMAGI 2000). Technically, these infrastructures are built on standards and methods developed by the OpenGIS Consortium (OGC) (http://www.opengis.org). Another OGC-based geodata infrastructure is currently developed at the University of Bremen within the framework of the GeoShare project.

GOALS OF THE GEOSHARE PROJECT

The overall goal of GeoShare is to support, promote and enable the utilization of geospatial data and spatial related knowledge with the help of information and communication technologies. The GeoShare project is funded through the EU Interreg IIIB North Sea Region Programme. Therefore the target area is the North Sea region, involving both urban and rural regions in four countries, namely Germany, the Netherlands, Norway and the United Kingdom (Fig. 1).

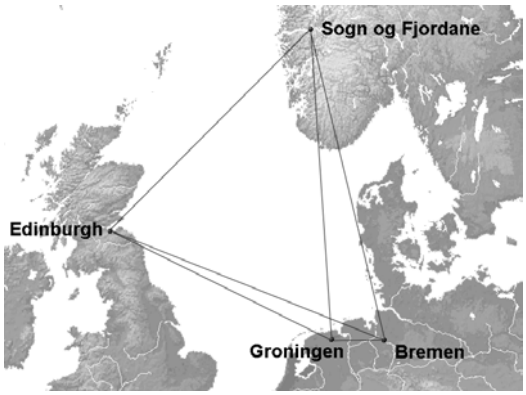

Fig. 1: The North Sea Region and the GeoShare partner cities

Within the framework of the GeoShare project, a technical infrastructure will be implemented. This infrastructure is designed to integrate the four partner regions in their attempts to improve the management and interpretation geospatial data and to give a significant added value to a number of already existing non-integrated systems and services. Against the background of the growing amount of spatial related knowledge in the North

Sea Region, GeoShare can be viewed as a pilot project to establish transnational regional information networks.

The result of GeoShare will be the provision of a framework, toolset and guidelines for sharing information through a GIS based system. This system will be validated through a set of web-based applications, each of which is focused on one of three thematic areas, namely

- spatial management,
- employment and education, and
- cultural/natural heritage and tourism.

In each of the four partner regions, multiple applications covering at least two thematic areas will be implemented.

The technical approach chosen to implement the GeoShare network builds on the experience of other projects and initiatives like the GDI-NRW and DataShare, a thematically related project that was funded by the Interreg-IIC North Sea Region Programme. In DataShare, the opportunities were explored that are embedded in web-based mapping technology to develop "geo-referenced data-sharing intranets" with the goal to facilitate integrated spatial planning. DataShare also investigated the major organizational and legal issues relevant to sharing data over distributed networks (e.g., the Internet and large Intranets), and to implementing prototypical data-sharing services. The GeoShare project builds on the key results of this previous project and moves on towards the development of high quality user-centric services within a broad range of application areas. GeoShare started in December 2002 with a total project duration of 3 years. This paper outlines the overall conceptual and technical approach to the geodata infrastructure, which will be implemented within the framework of the project.

THE GEOSHARE NETWORK

Recently, OGC standards and specifications started to play an important role in the design and implementation of geodata infrastructures. Since the foundation of the OGC initiative in 1994, most of the technical and conceptual challenges with regard to interoperable geospatial data and applications have been addressed. Based on a general conceptual approach (e.g. OGC 1999a), the OGC has issued implementation specifications for some of the most important technical topics (see http://www.opengis.org), while a number of related topics are still in the stage of discussion papers.

One major paradigm shift inherent in the OGC conceptual approach is the breakup of monolithic GIS applications into a number of modular geo-

services. These services are based on a common set of specifications and standardized interfaces that make them interchangeable and interoperable. The technical feasibility of these concepts and the resulting services-based infrastructures have been tested and proven by a number of projects and initiatives (e.g., the implementation testbed developed within the GDI-NRW). At the same time, OGC-conformant software is implemented by a growing number of GIS software providers, both in the commercial and the open source sector. As of May 2003, the OGC web pages (http://www.opengis.org) list more than 289 products registered as "OGC conformant" or "implementing OGC specifications and interfaces".

The transnational geodata infrastructure that will be implemented within the framework of the GeoShare project builds upon the developments described above. The architecture of this GeoShare *Network (GSN)* will be based on OGC specifications. Where possible, the individual geoservice modules that are part of the network will use standard OGC compatible software that is available through commercial and, preferably, non-commercial (i.e. open source) vendors. Where new or enhanced services are needed (like in the case of an intelligent search and retrieval service for resources in the network), innovative software developed by the project partners will be integrated.

The GeoShare Network will incorporate a number of different geoservices. All services will be made available to the GeoShare partners as well as to a number of organizations associated to the project. Within the project, a distinction will be made between basic and advanced services. *Basic services* are generic OGC compatible geoservices that are mainly used to store, access and display geodata in an OGC-compatible geodata infrastructure. These include *OGC Web Map Services (WMS)*, *OGC Web Feature Services (WFS)* and OGC compliant geodatabases. *Advanced services* are services that will be needed on top of the basic services to cover a set of application-specific tasks. Among others, these may include a *Coordinate Transformation Service*, a *Statistical Data Analysis Service*, an *OGC Sensor Collection Service (SCS)*, a *Service for Mobile Data Access*, and a *Web Gazetteer Service*. One particularly important advanced service is the *Intelligent Catalog Service (ICS)*, which will be the core of the GeoShare Network.

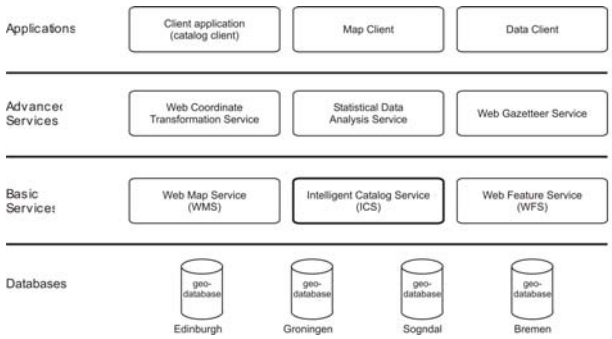

Fig. 2: Components of the GeoShare Network (overview)

In general, the project has to find a balance between what is technically possible (i.e., full-scale OGC-compatible service chains) and what is feasible given the organizatorial and technical constraints in the partner regions. Whether there is a large discrepancy between these extremes depends mainly on the willingness of the project partners to transform the already existing infrastructures and data-exchange policies to match the new paradigm. Experiences from previous projects and the preliminary analysis of infrastructures and policies in the partner regions indicate that in particular the organizatorial implementation of the new infrastructure could be problematic. Therefore, a phased implementation strategy was proposed which will ensure a smooth transition from the existing to the new infrastructure:

Phase I: A centralized catalog service is implemented as the cornerstone of the infrastructure. This catalog holds metadata descriptions of all applications registered within the network. In Phase I it will contain information about services, applications and information sources that are already in place within the project regions. The catalog service provides a user-friendly tool for the access to the resources in the network, as well as an interface that links the GeoShare network as a whole to other OGC compatible search tools. In Phase I, the applications hosted by the project partners have to be accessed through application-specific clients.

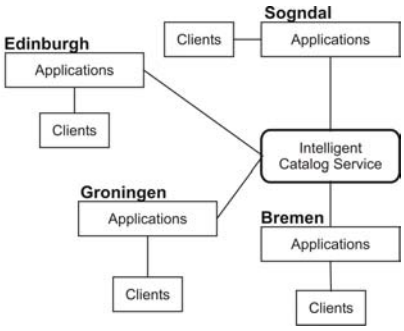

Fig. 3: GeoShare Network Phase I (schematic)

Phase II: In addition to the catalog service, a number of basic services forming the backbone of the network will be implemented and hosted at a centralized location. These services will include OGC compatible WMS and WFS. The services will be accessible to the project partners in a way that each partner can use them to configure application-specific web map services and web feature services. Through the respective OGC interfaces, these services will have access to data stored at the partner locations in OGC compatible geodatabases. If the need arises due to performance or data management requirements, additional basic services will be implemented at selected partner locations. Like the centralized basic services, these services will be registered and accessible through the centralized catalog service.

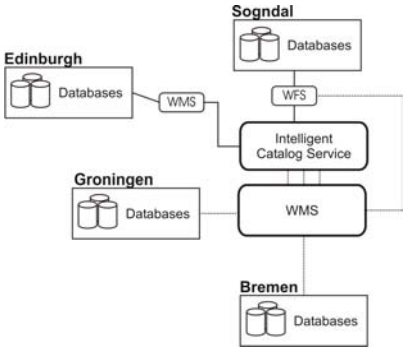

Fig. 4: GeoShare Network Phase II (schematic)

Phase III: Depending on the needs that arise through the implementation of advanced applications in the partner regions, additional advanced services will be developed. These services will be implemented and managed at the respective partner locations, but made available to the other project partners.

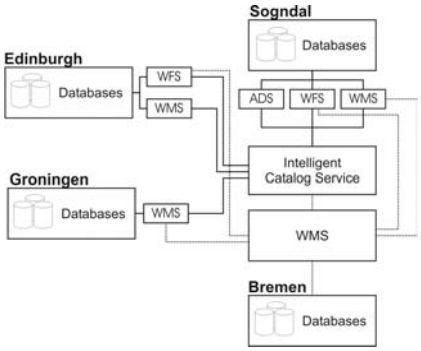

Fig. 5: GeoShare Network Phase III (schematic)

GEOSHARE SERVICES

As mentioned above, a number of basic services will shape the backbone of the GeoShare Network. The following is a short review of the basic characteristics of the basic services that are planned for project phases I and II:

Web Map Services (WMS): A Web Map Service produces map-renderings of geo-referenced data. It is important to note that in a WMS, a "map" is simply a visual representation of geodata that does not contain the data itself. A map may be rendered by the WMS in raster based image formats (PNG, GIF, JPEG) or as vector-based graphical elements in the Scaleable Vector Graphics (SVG) and Web Computer Graphics Metafile (WebCGM) formats. The Web Map Server Implementation Specification is an approved OGC specification and currently available in version 1.1.1 (OGC 2002a). So far, this specification has been implemented in a number of tools, both open source and commercial[1]. A standard WMS supports the `GetCapabilities` operation (to obtain service-level metadata), the `GetMap` operation (to obtain a rendered visual representation of the re-

[1] A list of OGC compatible tools is available on the OGC web pages (http://www.opengis.org)

quested data), and the `GetFeatureInfo` operation (to query attribute data specific to features displayed on the map).

An important feature of the WMS is its capability to create *composite maps*. Though the `GetMap` operation, a WMS can access a map on another WMS, allowing for the creation of a network if there are interconnected distributed WMS.

Web Feature Services (WFS): Like the WMS, the Web Feature Service implements an approved OGC implementation specification. The OGC Web Feature Service Implementation Specification is currently available in version 1.0.0 (OGC 2002b). A WFS provides an access to geospatial data stored in a geodatabase through stateless HTTP requests. The results of such requests are returned in GML format, an XML-based data exchange format for geospatial vector data (OGC 2003a). A WFS may be implemented as a basic read-only service. A *Basic WFS* applies the `GetCapabilities`, `DescribeFeatureType` and `GetFeature` operations to access service metadata, access feature metadata, and, most importantly, directly retrieve feature data. A WFS may also be set up as a *Transaction WFS*, which allows for two-way communication with the geodatabase through an `Transaction` and an (optional) `LockFeature` interface. In the first implementation steps of the GeoShare project, only Basic WFS will be implemented.

Geodatabases: As a pre-requisite to a meaningful interaction with WMS and WFS, the relevant geodata have to be stored in a geodatabase. State-of-the-art geodatabases overcome the (artificial) separation of geometric and attribute data and allow for the modeling of geographic objects in an object-relational schema. This data model is much closer to the OGC concept of geographic features (OGC 1999b) than most file-based data formats. Currently, a number of OGC compatible geodatabase tools are available, both in the open source and the commerical world (e.g., PostGres/PostGIS, ArcSDE, Oracle Spatial, etc.).

The Intelligent Catalog Service (ICS): The Intelligent Catalog Service will play an important and central role in the GeoShare Network. It is used to manage the metadata descriptions of services, data sets, and data collections in the network. The ICS will thus provide a detailed inventory of all registered resources, and queries to the ICS are the basis for the retrieval and selection of task-specific data and services.

The GeoShare ICS will integrate the evolving OGC specifications for catalog services (OGC 2002c) with BUSTER, a tool for intelligent information retrieval developed at the Center for Computing Technologies at the University of Bremen (Visser, Vögele et al. 2002). BUSTER, the *"**B**remen*

***U**niversity **S**emantic **T**ranslator for **E**nhanced **R**etrieval"*, is an information broker middleware that was prototypically built to support intelligent search and semantic data integration in distributed environments like the Semantic Web or, as in the case of GeoShare, a distributed geodata infrastructure. BUSTER is able to rank information resources based on their conceptual, spatial, and temporal relevance relative to a specific information request (Visser, Stuckenschmidt et al. 2002, Vögele and Schlieder 2003, Hübner 2003). BUSTER can also be used to integrate heterogeneous information through the resolution of structural, syntactical, and semantic heterogeneities.

The BUSTER module integrated in the GeoShare ICS supports queries of the type *concept @ location **in** time*. This means that in addition to the thematic relevance, the system is able to evaluate the spatial and the temporal relevance of resources that are registered with the catalog service. BUSTER uses reasoning algorithms based on formal representation languages to answer such a question in the thematic (conceptual) dimension ("Which concepts are subsumed by the concept specified in the query?"), the spatial dimension ("Which locations are spatially relevant to the specified geographic region?"), and the temporal dimension ("Which time intervals are relevant with respect to a given time period?").

GEOSHARE APPLICATIONS

In the following, we present two typical applications that will be implemented within the GeoShare project and integrated into the GeoShare Network. The first one is a *Water Directive Municipal Information Viewer* for the *Sogn og Fjordane Municipality* in Western Norway. In this municipality there is a need to provide citizens and local authorities with up-to-date information about issues related to environmental protection in general, and the quality of all surface- and groundwater resources that are used as drinking water sources in particular. This need arises as a response to national legislation which states that the citizen of Norway have the legal right to access information on local water quality in order to evaluate the quality of the services that are provided.

The core functionality of the planned application is to provide a map-based interface visualizing all sources of drinking water, as well as the related background data and environmental information. This includes access to up-to-date water quality information based on the analysis of regularly collected water samples. In addition, the application will provide tools to update and maintain a regional database of water quality indicators, and to

provide an interface that allows exchanging the data in a generalized form to the water quality database run by the national environmental protection agency.

Because the main purpose of the application is to inform the public about water quality related issues, the map interface has to be designed to meet the needs of end-users any skill level. This requires the development of a simple and intuitive user interface that allows the end-users to access information and data through a digital map with neither specific GIS skills nor knowledge of the terminology typically used by hydrologists, chemists and environmental engineers. Another challenge will be to enable the exchange of data and metadata between local, regional, and national data catalogs and geodatabases on various generalization levels. Here, problems of syntactic as well as semantic data interoperability will have to be solved.

From a technical point of view, most of the geospatial data that are relevant to this application are available in standard file-based data formats (ESRI shape files). Only a small fraction of the data is already integrated in a geodatabase based on the open source tools *PostgreSQL* and *PostGIS*. Eventually, all relevant geospatial data will be migrated to this geodatabase. For the online presentation of the data, an OGC compatible WMS based on the open source tool *UMN MapServer 3.6/3.7* is being implemented. The integration of the Municipal Information Viewer therefore skips Phase I in the GeoShare implementation plan (see above). The respective architecture will be compatible with Phase II, i.e., basic services (a geodataserver and a WMS) will be implemented locally and registered with the central catalogue service.

A very similar application is being built for the *Senator für Bau und Umwelt (SBU)* in Bremen, Germany. Like in *Sogn og Fjordane*, the intention of this application is to provide online and map-based access to data related to the implementation of the EU water framework directive (WFD). Among others, this directive regulates the way in which data pertaining to the quality of surface and ground waters have to be gathered, evaluated and reported. Apart from streamlining the reporting process within the national and EU administration, the directive calls for making the information available to an interested public in an appropriate way. The application will provide interactive access to visual renderings of a range of data related to surface water bodies in Bremen, including water quality measurements, specific biological quality indictors, the condition of river banks, and locations of waste water point sources.

A GENERIC GEOSHARE USE CASE

The applications described above show two typical examples of a number of applications in the GeoShare Network. From the analysis of these proposals we are able to extract a generic use-case for the GeoShare Network. Three main types of actors play a role in this use-case:

- information providers, i.e., organizations that produce, hold, and manage geospatial or geo-referenced data,
- information users, i.e., organizations or individuals that need (geo)data in support of specific tasks (e.g. information, decision making, planning, marketing), and
- service providers, i.e., organizations that maintain (web)services which are able to process the raw data provided by the information providers in a way to better match the specific needs of the information users.

In the initial phases of the GeoShare project, the role of the service providers will be taken over either by the project consortium itself through the establishment of a set of centralized basic services, or by the information providers through the implementation of local basic services. However, to develop a sustainable platform for the GeoShare Network, the responsibilities for common services may be handed over to specialized service providers during or after Phase III.

Through the GeoShare Network, information providers will be able to use a set of basic web services, for example to present geo-referenced information on the web. This includes the rendering of digital maps as well as object-based queries for the underlying attribute data. Applied to the first application described above this means that *Sogn of Fjordane Municipality* may use the central GeoShare web map server to set up a number of customized web map services to display information about the quality of surface water bodies, the locations of drinking water wells and sampling points, and other related data. Through the central web feature server, the information provider may set up protected web feature services to make the underlying geospatial data available for further processing to selected clients, for example other public organizations within *Sogn of Fjordane Municipality*.

All available resources in the network, i.e., data sets, data collections, WMS, and WFS will be registered with the GeoShare Intelligent Catalog Service. The registration points to a metadata description of the respective resource in XML/RDF format. Besides the standard ISO19115 compatible metadata, this *Comprehensive Source Description (CSD)* contains pointers

to the conceptual, spatial and temporal reference ontologies that were used to describe the resource. The ontologies used by the GeoShare Intelligent Catalog Service are formal and machine-readable representations of conceptual, spatial and temporal conceptualizations. They enable the ICS to "understand" the semantics of keywords, location references through place names, as well as time references through periods, and to use this knowledge for enhanced information retrieval.

For the information provider, the architecture of the CSD has the advantage that he can use a very sparse metadata set (i.e., using only those keywords, place names, and time periods that are most relevant) to describe his resources. In addition, this description can be specified using the information providers domain-specific terminology.

In the following, it is shown how *Sogn of Fjordane Municipality* uses a thematic ontology based on the environmental GEMET thesaurus, a spatial place name model based on the administrative subdivisions of Norway, and a temporal model based on memorable events and time periods in Western Norway.

On the other hand, the information user may apply the Intelligent Catalog Service to search for specific information in the GeoShare Network. Not only may he combine a query to search in the thematic, spatial and temporal dimension (*concept @ location* ***in*** *time*). He may also do so using thematic, spatial and temporal ontologies of choice. For example, a typical information request in the GeoShare thematic application area of cultural heritage and tourism could be: "Is there an information resource within the network that covers *fishing waters* in *Western Scandinavia* within the *last 2 years*.?" (see Fig. 6).

To specify this query in the ICS, the information user would resort to an ontology of recreational and touristic terms, a spatial model of tourist regions in Scandinavia, and a generic temporal model covering the last 10 years. The ICS will map the *query ontologies* to the *domain ontologies* used by the information providers for the metadata description of their resources. In the example given above, the tourist ontology will be mapped to an ontology of environmental terms, the spatial model of touristic regions to a spatial model of administrative subdivisions, and the generic temporal model to a temporal model of memorable events in Western Norway.

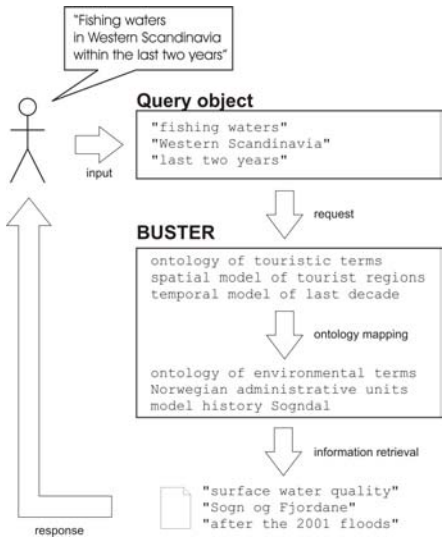

Fig. 6 Example of a query in the GeoShare ICS

Based on this mapping, the ICS is able to reason that a resource that was annotated for example with the keywords "*surface water quality*", the locational reference "*Sogn of Fjordane"*, and the period name *"after the 2001 floods"* is relevant to the information user's request.

DISCUSSION AND OUTLOOK

This paper outlines the approach taken to implement a transnational geodata infrastructure within GeoShare, a project funded through the EU Interreg IIIB programme. The GeoShare Network will support the exchange and usage of geodata within, as well as between the partner regions. Technically it will be based on OGC standards and specifications, using standard commercial and open source products where possible. A major innovative component will be the implementation of an Intelligent Catalog Service based on tools for semantic information retrieval and integration that were developed by GeoShare partners.

The GeoShare project is still in its starting phase and the planned network is currently in the stage of being implemented. Therefore the technical concept will have to be adjusted to problems encountered during the implementation of the network. However, through the phased implementation plan, we hope that the transition from the "old world" (of

plan, we hope that the transition from the "old world" (of stand-alone GIS applications) to the "new world" (of a services-based geodata infrastructure) will be a smooth one.

REFERENCES:

Bernard, L. (2002): *Experiences from an implementation testbed to set up a national SDI*. 5th AGILE Conference on Geographic Information Science, Palma de Mallorca.

Greve, K., L. Plümmer (2002): *Abschlusskommunique des D21-Kongreß Geoinformationswirtschaft 2002*, Bonn. Juni 2002.

Hübner, S. (2003): *Qualitative Abstraktion von Zeit für Annotation und Retrieval im Semantic Web. Dept. of Computer Science (FB 3)*. Bremen, Bremen University.

IMAGI (2000): *Konzeption eines effizienten Geodatenmanagements des Bundes, Interministerieller Ausschuss für Geoinformationswesen* (IMAGI).

OGC (1999a): *The OpenGIS Abstract Specification*, Open GIS Consortium.

OGC (1999b): *The OpenGIS Abstract Specification Topic 5: Features, Version 4*, OpenGIS Consortium.

OGC (2002a): *Web Map Server Interface Implementation Specification, V 1.1.1*, OpenGIS Consortium.

OGC (2002b): *Web Feature Server Interface Implementation Specification, Version 1.0.0*, OpenGIS Consortium. 2002.

OGC (2002c): *OpenGIS® Catalog Services Implementation Specification, Rev: 1.1.1*, OpenGIS Consortium.

OGC (2003a): *Geography Markup Language (GML) Implementation Specification, Version 3.00*. S. Cox, P. Daisey, R. Lake, C. Portele and A. Whiteside, OpenGIS Consortium.

Visser, U., H. Stuckenschmidt, G. Schuster, T. Vögele (2002): *Ontologies for Geographic Information Processing*. Computers & Geosciences 28: 103-117.

Visser, U., T. Vögele, C. Schlieder (2002): *Spatio-Terminological Information Retrieval using the BUSTER System*. In: Proceedings of Environmental Informatics 2002, Vienna, Austria.

Vögele, T., C. Schlieder (2003): *Spatially-Aware Information Retrieval with Place Names*. In: Proceedings of the FLAIRS 2003, AAAI Press.