GEOLOGIJA 46/2, 343–348, Ljubljana 2003

The Architecture of the Georgia Basin Digital Library: Using geoscientific knowledge in sustainable development

B. BRODARIC[1], M. JOURNEAY[2], S. TALWAR[2,3], R. HARRAP[4], J. van ULDEN[2], R. GRANT[5] & S.DENNY[2]

[1]Geological Survey of Canada, 234B-615 Booth St., Ottawa, ON K1A 0E9
[2]Geological Survey of Canada, 605 Robson Street, Suite 101, Vancouver, BC V6B 5J3
[3]University of B.C., Dept. of Geography, 1984 West Mall, Vancouver, BC V6T 1Z2
[4]Queens University, Dept. of Geological Sciences, Miller Hall, Kingston ON, K7L 3N6
[5]University of B.C., SDRI, 1924 West Mall, UBC Vancouver, BC Canada V6T 1Z2
{bmjourne,stalwart,jvanulde,rgrant,sdenny}@NRCan.gc.ca; harrap@geol.queensu.ca

Key words: Digital Library, architecture, geoscientific knowledge, GIS

Abstract

To address societal issues such as sustainable development, geoscience knowledge must be transformed from its conventional packaging. A holistic approach to this transformation requires that (1) geo-information be used to develop indicators such as mineral potential, material availability, groundwater capacity, etc., and that (2) the indicators are combined with socio-economic factors to guide generation of future scenarios for land use, population, urban growth, etc. In this paper we discuss the architecture of a prototype system designed to support these two activities. The system consists of a geoscientific data and knowledge repository that is loosely linked with a future scenario modeling tool (Quest™). The repository supplies transformed information to the tool, and also provides explanations for the input variables and output results. An example implementation, Georgia Basin Digital Library, is also presented.

Introduction

It is commonly understood by geoscientists that geoscience information has a vital role to play in how significant issues, such as climate change, sustainability, biodiversity, natural hazard prevention and mitigation, are to be addressed. Though the importance of geoscientific information may be self-evident to geoscientists, the information itself is nevertheless generally overlooked by non-geoscientists dealing with those issues. Improving accessibility to geoscientific information through database construction and web delivery raises the visibility and availability of geoscientific information, but may not increase its use in non-geologic domains where the information is not well understood, largely because the information is complex and primarily aimed at geoscientists rather than at biologists, environmental scientists, etc. Geoscience information must therefore be transformed from its conventional packaging, typically as a map, to modes specifically usable by other disciplines prior to its delivery to those disciplines. This transformation may involve the development of geoscientific indicators that reduce significant geological situations into a single region-specific value for prediction or assessment purposes, such as estimates of mineral resource potential, building material availability, groundwater capacity, etc. It may further involve the incorporation of such geoscientific indicators into broader models that integrate natural resource indicators with socio-economic factors for purposes of generating region-specific predictions for land use, population growth, urban expansion, etc.

In this paper we focus on the latter approach and discuss the architecture of an internet-based system that integrates geological information into natural resource-socio-economic models. The system was developed within the Georgia Basin Futures Project (www.basinfutures.net/), an academic research project carried out at the University of British Columbia (UBC) and the Geological Survey of Canada (GSC). The ultimate goal of the project and resulting system is to better inform decision-making in sustainable development by allowing future scenarios to be developed and inspected. The main notion here is firstly, that using geoscientific information as an additional input will improve the accuracy and utility of future scenarios, and secondly, that exploring the nature and consequences of future scenarios will ultimately lead to better decisions. To achieve these two goals we develop a system that links two main modules: (1) an information and knowledge repository (i.e. *ontology + data*), called the Georgia Basin Digital Library (Journeay, et al., 2000), that allows relevant information and related knowledge to be stored, accessed and explored, and (2) a modeling tool (Quest™), from UBC, that uses the repository information to generate future scenarios. These linked modules constitute a simple decision support system for sustainability.

The Georgia Basin digital library

The knowledge and information repository is structured as a web-based digital library that contains 64 significant geospatial information layers such as land use, transportation, various natural resource layers, etc. It also contains a network of *concepts*, organized as *ontologies* and described using *stories*, to explain what role the layers and their contents play in the sustainable development of the region. So, the repository not only holds information but also possesses knowledge elements that explain what the information means and how it applies to sustainable development. The repository can thus be viewed as also supplying a knowledge layer to the stored information. Because initial implementation of the repository is focused on the southwestern region of British Columbia, the repository is called the Georgia Basin Digital Library (GBDL).

Coupled to the repository is a web-based user interface that enables repository contents to be viewed and explored. The interface also connects the repository contents to other relevant sources of information and knowledge, such as publications in other digital libraries, and news stories from web-based news agencies. It further allows individuals and groups to input their own geospatial sites and descriptions, to foster the input of local knowledge and thereby stimulate dialog on issues of sustainability in the context of a community or region. This user interface is called the Georgia Basin Explorer (GBX). The repository is coupled with the user interface via web-enabled functions. As shown in Figure 1, the system architecture has 3 tiers: information, web-services and presentation.

The information tier

There are three types of information sources that together comprise the information and knowledge repository: metadata, geospatial layers, and a knowledge-database. The metadata component stores in-

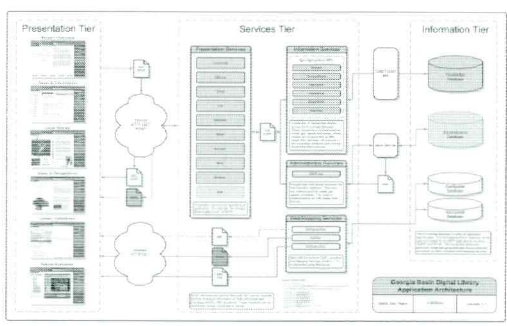

Figure 1. The three tier technical architecture of GBDL.

formation about users, projects, and information content. The 64 geospatial layers represent different geospatial themes for the Georgia Basin and are stored as ESRI shape files. The knowledge-database maintains the ontologies in a relational database (SQL Server): it consists of concepts, their relationships to each other, and to the metadata and geospatial data. The Figure 2 depicts the main elements of the logical schema used by the knowledge-database, and illustrates how geologic information can be stored in the design (c.f. Brodaric & Hastings, 2002; also Brodaric & Gahegan, 2002).

The Service Tier

The Service Tier provides a suite of web-based functions that manipulate the contents of the Information Tier. Spatial data is exposed via the WMS service specified by the OpenGis Consortium (www.opengis.org), metadata about spatial layers is exposed via the Z39.50 protocol, and project metadata and the knowledge-database are exposed through a custom-built web service. The service tier thus serves as a bridge from the repository to external web clients who can be either persons or software agents. It also serves as an internal bridge between the repository, metadata and spatial data for operations internal to GBDL such as building and updating the repository. The custom web interface is divided into two sets: low-level foundation services operating on the knowledge-base, spatial data (WMS) and metadata (Z39.50), and high-level presentation services that perform dedicated procedures for specific interface components. The services are currently accessed via http, but we are migrating these to SOAP/WDSL platforms. XML encoding standards are used to pass requests and results between tiers (Figure 3).

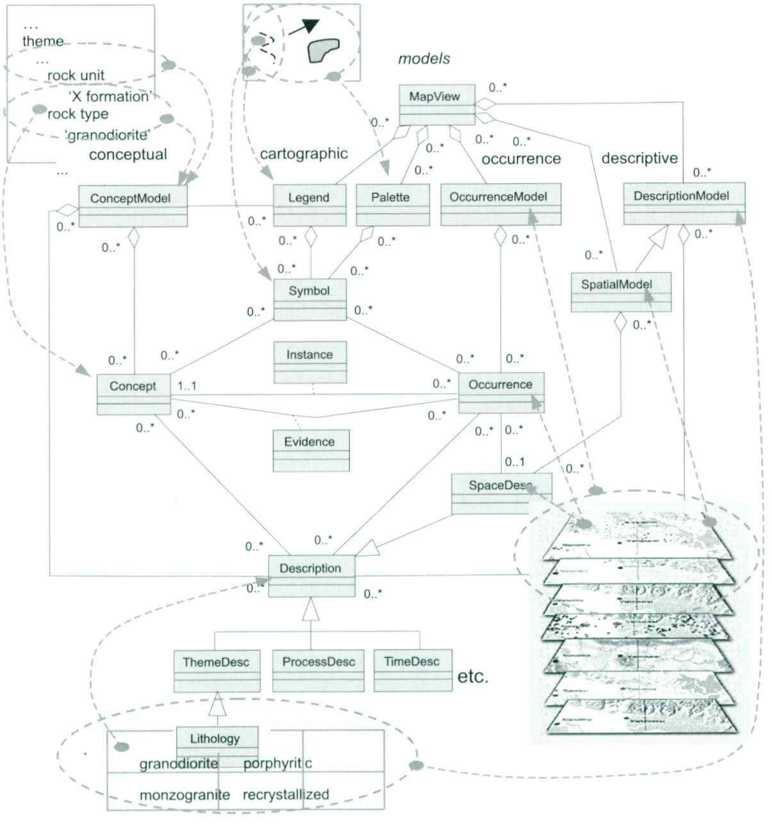

Figure 2. Logical schema for the knowledge-database (c.f. Brodaric & Hastings, 2002).

```
<?xml version="1.0" encoding="UTF-8" ?>
- <wsdl:definitions targetNamespace="http://geosemantica_api" xmlns:wsdl="http://schemas.xmlsoap.org/wsdl/"
    xmlns:xsd="http://www.w3.org/2001/XMLSchema" xmlns:wsdlsoap="http://schemas.xmlsoap.org/wsdl/soap/"
    xmlns:intf="http://geosemantica_api" xmlns:impl="http://geosemantica_api-impl" xmlns:SOAP-
    ENC="http://schemas.xmlsoap.org/soap/encoding/" xmlns="http://schemas.xmlsoap.org/wsdl/">
  + <types>
  + <wsdl:message name="getConceptResponse">
  + <wsdl:message name="createConceptResponse">
  + <wsdl:message name="getConceptRequest">
  + <wsdl:message name="deleteConceptResponse">
  + <wsdl:message name="createConceptRequest">
  + <wsdl:message name="getConceptsRequest">
    <wsdl:message name="CFCInvocationException" />
  + <wsdl:message name="getConceptsResponse">
  + <wsdl:message name="deleteConceptRequest">
  + <wsdl:portType name="concept">
  + <wsdl:binding name="concept.cfcSoapBinding" type="intf:concept">
  + <wsdl:service name="conceptService">
  </wsdl:definitions>
```


Figure 3. Example requests to the web services tier, encoded in XML.

The Presentation Tier

The presentation tier (GBX) provides a user interface to the information and knowledge repository—it acts as a client to the services tier, exclusively using the services tier to access the library contents. GBX aims to build awareness and understanding about sustainability amongst municipal and scientific communities in the Georgia Basin by directly engaging those communities using the world-wide-web. GBX operates by using the services tier to access the information and knowledge repository. It has five components: (1) *News and Information*, which connects specific concepts from the knowledge-database with topically relevant web-based news stories and other pages; (2) *Library Collections*, which uses specific concepts to search web-based catalogs for maps, images and reports and displays them for viewing; (3) *Local Stories*, which allows users to annotate geospatial layers with personal geospatial sites and related stories that convey local knowledge about issues of sustainability in the community or region; (4) *Ideas and Perspectives*, which enables concepts and their story-like descriptions (from the knowledge-database) and related geospatial layers to be explored interactively; and (5) *Future Scenarios*, which connects with the Quest modeling tool for generating future scenarios for the region. Particularly significant is the *Ideas and Perspectives* component, as it presents community, academic and NGO (non-govn't organization) perspectives on sustainability, each organized as a separate ontology in the knowledge-database. Figure 4 displays the *Ideas & Perspectives* component, which enables the perspectives (ontologies) to be explored by browsing a semantic network of concepts (top left) triggering the display of related maps (bottom left) and story-like documentation (right).

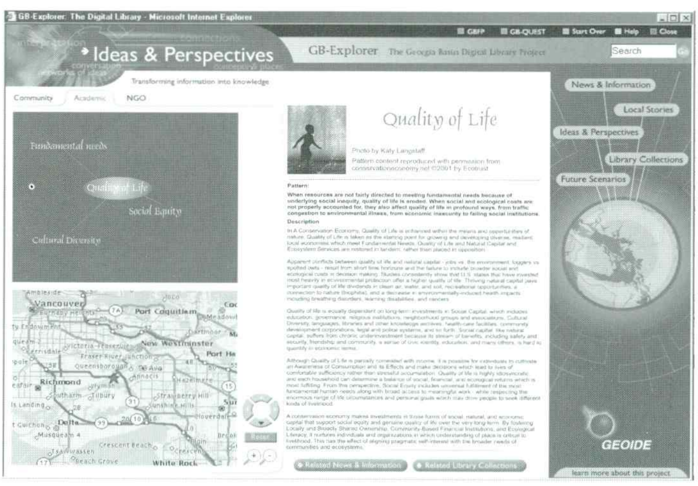

Figure 4. Exploring sustainable development concepts in GBDL.

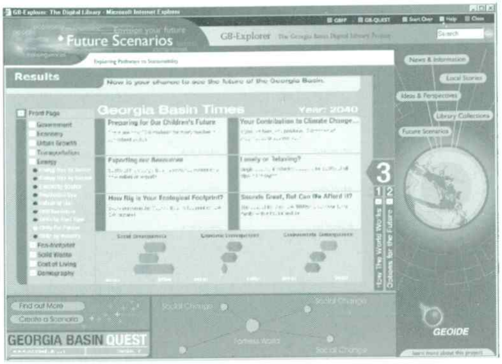

Figure 5. Explaining the results of a Quest future scenario in GBDL.

The quest scenario modeler

Apart from its representation and presentation roles, GBDL also provides dynamic connection to the Quest scenario modeling environment, in which future scenarios, such as land use, economic and demographic projections, are constructed. Quest is a computer simulation that enables people from all walks of life to construct alternative futures for a region and view the trade-offs and consequences of their choices. The scenarios are controlled by input from the user and are driven by an integrated suite of system models that reflect the expert knowledge of more than 35 researchers in the fields of environmental, social and economic health. In this modeling procedure GBDL is designed to serve as an information repository, supplying information to Quest for modeling, and as an explanatory tool, allowing the user to browse input variables and output results, thus helping build a context for understanding. The results of a Quest scenario are uploaded to the GBDL scenario library (Figure 5), where the various input parameters and modeling assumptions can be reviewed and further explored in the *Ideas & Perspective* module (described above). At the time of this paper, the dynamic interaction between the GBDL and Quest is still under development, with positive initial results.

Of broader significance is the fact that geoscientific knowledge is utilized as an input theme and therefore as part of the modeling system, clearly demonstrating the relevance of geoscience information in the context of sustainability planning and decision making. In our work to date we have prototyped a sub-model that integrates the potential socio-economic impacts of earthquakes into the Quest modeling framework; we are currently working on developing an integrated surface/groundwater sub-model to help address issues of sustainable yield and aquifer vulnerability in the region.

Conclusions

The GBDL concept has been prototyped in two geographic regions: Bowen Island and the Georgia Basin (http://georgiabasin.info/). The Bowen Island prototype is most mature to date, as it has been used extensively in a grade 10 experiential learning environment where students were asked to develop their own *Local Stories*. GBDL is just recently coming on-line as part of a larger research project, the Georgia Basin Futures Project (http://www.basinfutures.net/). Our experiences from this project indicate that the combination of knowledge representation techniques, modeling tools, geoscience information and the world-wide-web is practical, informative and useful. We anticipate developing these further, specifically by introducing more geoscientific information into the modeling component, tighter coupling of the modeling and knowledge repository components, and applying the system in different geographic regions.

References

Brodaric, B. & Gahegan, M. 2002: Distinguishing instances and evidence of geographical concepts for geospatial database design. In: Geographic Information Science-2nd Int'l Conference, GIScience 2002, M. Egenhofer and D. Mark (eds.), LNCS 2478, Springer, 22-37, Berlin.

Brodaric, B. & Hastings, J. 2002: An object model for geologic map information. In: Richardson, D., van Oosterom, P. (Eds.), Advances in Spatial Data Handling, 10th International Symposium on Spatial Data Handling, Springer, 55-68, New York.

Journeay, M., Robinson, J., Talwar, S., Walsh, M., Biggs, D., McNaney, K., Kay, B., Brodaric, B. & Harrap, R. 2000: The Georgia Basin Digital Library: Infrastructure for a Sustainable Future. Proceedings, GeoCanada 2000, May 29-June 2, Calgary.