OPERATIONAL GIS USER-INTERFACE FOR HYBRID GEO-DATA BASED ON DYNAMIC DATA RETRIEVAL

Charlotte Steinmeier
Swiss Federal Research Institute WSL,
CH-8903 Birmensdorf
Switzerland
steinmeier@wsl.ch

Working Group IV/1

KEY WORDS: GIS, User Interfaces, Meta-data, Spatial data

ABSTRACT

Geo-information nowadays encompasses data of different types and origins recorded at different time periods using a variety of methods. To exploit the full potential of this data and to use it in a synergistic fashion in image analysis for example, new techniques for data management and handling have to be developed. This paper is based on a process oriented data model, using a relational database management system. The database includes the varying data types, such as spatial data, attribute data and meta-data. The purpose of this study was to develop an operational interface for non-GIS, non-database specialists to facilitate them the usage of the data. The interface automatically creates small dialogs that offer the user the results of a database query. Based on the users' selection in the dialog new queries are composed and sent to the database. This dynamic data retrieval, in combination with the process oriented database, allows the user to find all related data of a selected spatial object and presents it, if desired, to the user. The automated creation of complex query statements enables the 'once in a while' GIS user to work with a complex database without any knowledge of the database structure.

KURZFASSUNG

Unter dem Begriff Geo-Informationen versteht man heutzutage Daten unterschiedlichen Typs und unterschiedlicher Herkunft, die zu verschiedenen Zeiten mit unterschiedlichen Methoden erfaßt wurden. Um das volle Potential aller Daten auszunutzen und um sie zum Beispiel für die Bildverarbeitung synergetisch zu gebrauchen, müssen neue Verfahren zur Datenverwaltung und zur Datenbearbeitung entwickelt werden. Die hier vorgestellte Arbeit basiert auf einem prozessorientierten Datenmodell, das ein relationales Datenbank Management System benutzt. Die Datenbank beinhaltet verschiedene Datentypen, wie z.B. räumliche Daten, Attributdaten und Metadaten. Der Zweck dieser Untersuchung war ein operationelles Interface für Nicht-GIS- oder Datenbank-Spezialisten zu entwickeln, um ihnen den Gebrauch der Daten zu erleichtern. Das Interface erstellt automatisch kleine Dialoge, die dem Nutzer die Ergebnisse einer Datenbankabfrage darstellen. Entsprechend der vom Nutzer getroffenen Auswahl im Dialog werden neue Abfragen zusammengestellt und zur Datenbank geschickt. Diese dynamische Datenabfrage, kombiniert mit der prozessorientierten Datenbank , stellt dem Nutzer alle Daten, die einen Bezug zum selektierten räumlichen Objekt haben zur Verfügung und präsentiert sie, falls erwünscht. Die automatische Erstellung der komplizierten Abfrage-Statements ermöglicht dem Gelegenheits-GIS-Nutzer das Arbeiten mit einer komplexen Datenbank, ohne Kenntnisse über Datenbankstrukturen.

1 INTRODUCTION

During the last two decades, spatial data has increasingly become available in digital form. This data is either scanned from analogue data or directly recorded with digital mapping tools. Digital geo-data can comprise very different data types such as vector data, point data, images or attributes. Its storage is usually specific for the soft- and hardware used, thus restricting its dissemination. Furthermore the retrieval and presentation of such data usually requires specialists who are able to work with Geographic Information Systems (GIS).

In recent years, research in the field of Geographic Information Systems has expanded in many different subtopics like spatial modeling, data semantics or data quality (Muller 1993). More and more data with a spatial context has become available (Flewelling 1999) and many subtopics have become a research field in their own right. Nowadays GIS applications are used in very different domains of interest, ranging from their classic use in environmental sciences, to real estate markets, to the management of power line networks, to logistics (Parrott 1991, Bennett 1997). The very variety of these fields of application and the ever increasing amount of available digital data create also new demands on the GIS systems. They should be able to handle, for example, different data types, very large data sets and temporal components. Moreover the GIS should allow data dissemination and user access (Worboys 1995).

To date, many specific problems have been discussed in research papers but there is little information available about the user's acceptance and usage of these systems. The described graphical user interfaces (GUI) deal mostly with static databases and do not include the extraction and presentation of meta-data. (Weinschenk 1997). Flewelling (1993) states that the task of *finding* the right data is different from that of *using* the data in scientific analyses and therefore requires different supporting tools. New techniques for data managment and data handling need to be developed in order to exploit the full potential of all this data and to use it for example in a synergistic fashion in image analysis. The integration of hybrid data (e.g. vectors, attributes, documents, sounds) in combination with a temporal component requires a new database design preferably within one single database management system (DBMS). The possibility to retrieve data from this system should be easy to use and allow a user who is neither a GIS specialist nor a database specialist to do it and look at it as if looking at paper printouts. Without an assisting user-interface to the database the data will be archived, accessible only to some experts and but will hardly ever be used by the interested geo-data user.

The work described here is part of a larger project conducted in collaboration with the Swiss Agency for the Environment, Forests and Landscape. The content of the project is the design and implementation of a database for the storage and management of data for the environmental protection of mires, riverine forests, semi-arid meadows, etc. (Grünig 1992). Besides the spatial data representing the geographic location and extent of the investigated objects, much more additional data has to be archived. The database should also comprise base documents such as images, documents containing the criteria for selecting protection areas and government documents which constitute the legal base for the protection of these areas. The objective of the present study was to develop an operational user interface which is easy to use for non-GIS- or database experts. This graphical user interface should enable the user to retrieve the desired spatial data and information connected with selected objects such as aerial images or administrative documents.

2 OUTLINE

Well known software components were chosen in order to offer easy access to a wide variety of users with different hardware systems. All the different types of data, such as spatial objects, attribute data and a variety of related meta-data are archived in a single Oracle database. The handling of the spatial data is managed by the Spatial Data Base Engine (SDE) of ESRI (ESRI, 1998). The database design of the project has to take into consideration the hybrid character of the data and must be able to store the complete selection process of the protection areas. A process-oriented data model was therefore introduced. It is described separately in Baumberger and Hägeli (2000). Such a complex database contains much information but in most cases only some parts of the data are of interest to the users. For example, some users will only be interested in the meta-data such as the soft-or hardware components used for data collection, others in the spatial information whilst others in turn will to retrieve the lineage.

In order to realize an efficient data retrieval these different perspectives of the users to the archived data had to be taken into account. The database model for the presented interface permits three different 'data access routes':

- the 'geographic' approach – the user first selects spatial objects from a map or from a list and obtains then any related information from other tables by joining the selection with the meta-data.
- the 'meta-data' approach –the user selects the data according to the type of meta-data, such as fix-points or documents.
- the 'process oriented' approach – the user retrieves the desired data by tracing the process chain.

The dynamic data retrieval described here is based upon the 'geographical' approach. In order to develop interfaces which are easy to use it is necessary to develop different interfaces for the specific needs of the users. Only the geographic approach will be presented and discussed in this article.

2.1 User Group

An interface should facilitate the communication between the user and a computer. In order to develop a functional interface, it is necessary to focus on a special user group. Therefore the starting point of the project consisted of a preliminary survey of the potential users with their specific requirements and technical abilities. Three groups could be identified: first, the *expert,* who has worked already for several years with GIS data and DBMS systems; second, the *every once in a while user,* who knows the basics of GIS – systems and who does not have the time and leisure to learn it intensively; and finally, the *employee*, who has to retrieve special information occasionally on request but who has never worked with digital spatial data. It was decided to develop the interface for the *every once a while user*. This group forms the majority and the interface would enable them to work very efficiently. The group mainly consists of scientists from other disciplines such as biologists or forest engineers who want to archive their data securely and efficiently and whose main interest is the analysis of the data in combination with geo-data. Within the group there are also some members who want to take a quick and easy look at the available spatial data and meta-data. They often have to retrieve data that is connected in different ways to the spatial data, for example, the law on which the inventory is based, or an aerial image showing the positions of fix-points.

2.2 Requirements

The design of a user-specific operational interface requires a good knowledge of the users' needs and problems. In the current research the specific requirements for the selected user group were gathered. An analysis of the survey revealed that the following should be taken into account:

- the data sets should be easily available for different users at different locations
- the GIS system should be available for the most common platforms (UNIX, PC, MAC)
- the GIS system should be fairly common among the users
- the GIS system should be cheap in case of unfrequent use
- the interface design should reflect the structure of the data archived in the ring binders
- the interface should lead the user directly to the presentation of the selected data
- the interface should be able to handle additional inventories without reprogramming it
- the interface should, as far as possible, be similar for all different inventories to ensure the easy usage
- the interface should also handle any other software to look, for example, at images or text documents
- the interface should allow the more experienced user to work with the GIS system without any restrictions.

For the GIS system, the choice was made in favor of ESRI Arcview (ESRI, 2000). Its script language *Avenue* allows not only the implementation of a very specific interface but also the usage of the original functionality of Arcview.

3 PROCESS ORIENTED DATABASE

The database archive itself is not static but expandable. Additional data sets can be added as well as additional data structures. This is necessary since more and more analogue inventories will be digitized and transferred to the database of national protection areas. In addition, it is important that corrections to existing data can be made and that field data from various monitoring schemes can be added. The possibility to add data to the database at the same time as to retrieve it for operational use requires dynamic data retrieval. This means more programming work at first but in the long run it is worth the effort since the dynamically programmed software modules can be reused for new projects.

Several steps are necessary to set up an environmental protection area. These steps may start, for instance, with a scientific analysis of a special area which reveals a national interest in protecting this region. A governmental decision will follow to initiate the selection process. For the next step basic documents such as aerial images and maps are used for digitizing the potential boundaries. Subsequently, these potential areas with the according boundaries are verified by field work and afterwards presented to the actual owners and local authorities for survey. Next, if necessary, the boundary is readjusted and finally, a legal document published by the government, determines the exact position and extent of the protection areas.

These clearly identifiable steps of delineating a protection area form the base for the data model. Each individual step is considered as a process which is related to a certain type of data and corresponding meta-data. The main components of the data model are the process types which are classes of similar process steps, and process instances which are the actual occurrences of process steps. Data resulting from a process step is associated with the process instance. The interrelationships between the different process types are modeled by associating a child process type with its parent types. The linking of the process types was essentially defined by a directed graph showing the lineage for a given node of the graph. The complete history of all entities stored within the database can be tracked by traversing the graph of processes. Figure 1 shows an example for a graph of linked process types.

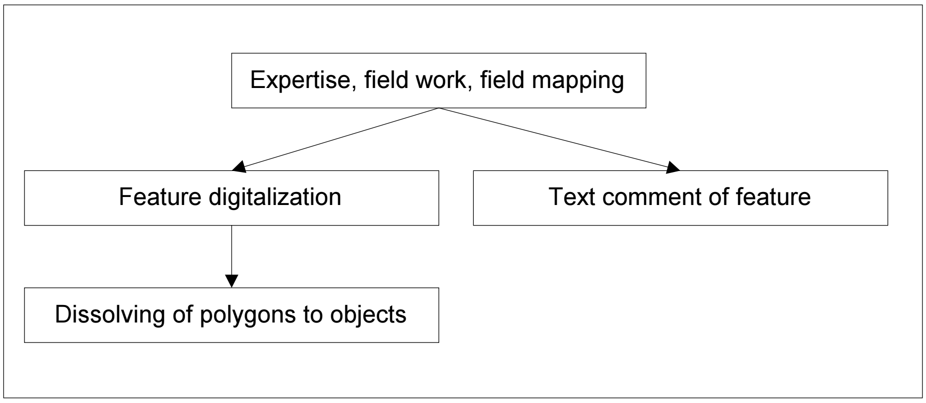

Figure 1: Example of directed graph for process type lineage

4 REALIZATION

A powerful GIS system requires specialists for effective usage. A lot of buttons and popup menus are available and permit fast and easy handling of spatial geo-data. But, as a rule, the *every once a while user* is irritated by all these software features. Much expensive time is required to do even simple things like presenting available data. In order to facilitate the handling of the GIS it helps already, to set only those buttons visible which facilitate the users' needs at a specific operating step. Self-evidently, this denotes a fairly strong restriction in the available functionality of the GIS but in return, it allows a guiding of the user. The pre-selection of the user group determines the usual operating cycles. In the presented case this means that the user needs a connection to the actual database, a window for the presentation of the spatial entities with the according help-dialogs and, often at the end of the working cycle, a utility to save or to print the selected objects together with their related information.

4.1 Dynamic Data Retrieval

The spatial approach forms the key criteria for the interface design. It is assumed that the users are mainly interested in the information about some or all spatial objects of a selected inventory. As mentioned above not only the attribute- and meta-data is stored in the Oracle database but also the spatial data. The spatial data of an inventory corresponds to all intermediate steps of boundary extents during the protection procedure. The handling of the spatial data is done with the Spatial Database Engine (SDE). Due to the data storage inside a relational database management system (RDBMS) the different intermediate steps of the protection areas' extents are only distinguishable by additional specifications from joined attribute- or process tables. In order to retrieve only some specific records of spatial data, it is necessary to create a SDE-query statement. The syntax of a SDE query resembles strongly a regular structured query language (SQL) statement. Arcview assists the expert user quite well using little dialogs but nevertheless it is absolutely necessary to have at least some knowledge about SQL. One of the most important requirements of the interface is therefore to free the user of the formulation of such query statements.

The key criteria developing an easy to use retrieval program is that all the information should be stored within the database, such as dialog titles, descriptive headers, help text for buttons and so on. Furthermore, it must include a language key since Switzerland harbors four official languages and often English is included as well.

The interface is newly created every time that it is started. This means that every dialog, button, icon, view window etc. is created as and when it is needed and deleted when the interface is exited. This ensures that only the actual data of the database is retrieved even if it is rather static data like state boundaries or digital maps. And it enables little updates or customizations to be made without the reprogramming of the interface. Following the startup of the interface the user is asked to enter the password for the database connection. Subsequently, the user has to choose one of the languages, popping up as check-button in a little dialog within the interface. Following this, a map of Switzerland with the state – and community borders appears in a view window and thereafter, the user has to select the inventory of interest.

The query statement itself is a rather complex query, since the access to specific spatial data records is possible *only* by analyzing the according meta data. Furthermore, the records are number-coded and must therefore be transformed with the help of according lookup tables into text information of the chosen language. Figure 2 shows an example of such a little dialog. In this example, the query searches first for the desired inventory (here: WZVV) and gets back six different surveys. Another query is formulated automatically with the help of the information of the previous query. It searches each of these surveys for the existence of different precision levels. (e.g. objects and subdivisions of objects). With the help of the process table it is possible to get the right records which then can be further analyzed or directly presented within the view. The guidance of the user is done by making these dialogs 'modal' - that is nothing else can be done before the selection of the spatial data is completed and confirmed.

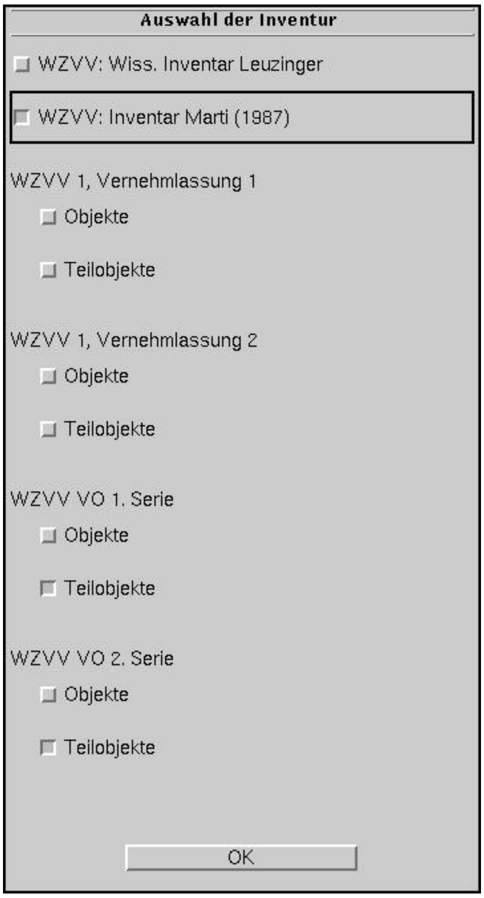

Figure 2: example of a dynamically created dialog based on an automatic creation of a query statement.

4.2 Dynamic GUI

As soon as the presentation of the retrieved data within a view is completed, additional control buttons became visible. These buttons and tools help the user to edit the legend for example, or to select objects within a selected state or community (overlay-function) or to select a subset of the retrieved objects. To get text information to the selected objects requires again a complex data query since these info-files are not a process step but data (LOB-files) which are affiliated to objects or subdivided objects.

The user is always guided by 'assistance dialogs'. They are created automatically until the user has chosen one or more objects which the information is required. The according query requests are generated automatically on the basis of the user's choice in the previous dialogs, i.e. the results of the query define the next query statement. If there are several sub-choices possible, a little interaction dialog is created to enable the user to lead the query in the desired direction.

One of the requirements listed above was to show all available information related to an object. This needn't be directly related such as object attributes but also documents with only a 'loose' connection, for example a legal text document or an image of the area. The only opportunity to obtain all the information not directly connected to an object is to analyze the process chain. The directed graphs of the process oriented data model allow the user to find all previous process steps of a given object. A detailed description of the lineage and its visualization is given in Brändli, (2000). The process-chaining returns all process types of the inventory investigated and from these, the corresponding process numbers of the selected object can be retrieved. According to the process type, different buttons became visible. The according information is either downloaded directly by a mouse-click onto such a button or, if several documents exist regarding the same process type, a small dialog opens with check-buttons showing the available options. The selected document will be downloaded to the file system, since it is unstructured information (text files, images, references). The GUI will then automatically load and open a PDF-reader for the text files or an available viewer for images.

5 CONCLUSION

More and more geo-information is available but often it is archived in large complex databases where experts are needed to retrieve the data. The emphasis of the presented work was to facilitate the access to geo-data. A user-interface was developed to enable people who are neither GIS specialists nor database specialists to use a database with hybrid geo-information. The user-interface as well as the database design are not yet tested extensively. However, first tests with the envisaged customers give encouraging results and show the promising potential of the presented approach.

The process oriented database model allows the storage of spatial data as well as meta-data and attributes within a single relational database management system. The usage of meta-data for the selection of spatial information makes the retrieval queries rather complex, but it offers the possibility to add new data sets as well as new data structures without any reprogramming of the interface modules. The user is guided through the database by automatically created SQL-queries. These are defined by the users' selections in the dialogs which, in turn, are based on the results of the previous queries. The presented work allows the user to access and work with hybrid data within a spatial context. The next step will be the transform into a web-based application. This offers the possibility to make the inventories of the protection areas readily available and accessible so that they can be used for a large variety of applications.

6 REFERENCES

Baumberger, N. and Hägeli, M., 2000. Using Meta data in Multi Step Preprocessing and long term monitoring. Submitted Paper, GIS Work Session UN/ECE, Conference of European Statisticians, April 10-12, Neuchatel, Switzerland.

Bennett, D. A., 1997. A framework for the integration of geographical information systems and model base management. In *Int. J. Geographical Information Systems,* Vol. 11, 337-357.

Brändli, M.,2000. Effective compilation and browsing of geographical meta-data using common interoperability tools. Submitted Paper, GIS Work Session UN/ECE, Conference of European Statisticians, April 10-12, Neuchatel, Switzerland.

ESRI, 1998. Spatial Database Engine. ESRI White Paper, Environmental Systems Research Institute, Inc., Redlands, CA.

ESRI, 2000. Homepage ArcView: http://www.esri.com/software/arcview/index.html

Flewelling, D. M., and M. J. Egenhofer, 1999. Using digital spatial archives effectively. In *Int. J. Geographical Information Systems,* Vol. 13, No. 1.

Flewelling, D. M., and M. J. Egenhofer, 1993. Formalizing importance: parameters for settlement selection. In *Proceedings of 11th International conference on Automated Cartography*, edited by R. McMaster (Minneapolis, MN: American Congress on Surveying and Mapping), pp. 167-175.

Grünig, A., 1992. Mires and man: Mire conservation in a densely populated country – the Swiss experience. *Excursion Guide and Symposium Proceedings of the 5th Field Symposium of the International Mire Conservation group (IMCG), Switzerland*. Swiss Federal Institute for Forest, Snow and Landscape Research, Birmensdorf, Switzerland.

Muller, J.-C., 1993. Latest developments in GIS / LAN. In *Int. J. Geographical Information Systems*, Vol. 7, No. 4, 293-303.

Parrott, R., and Stutz, F. P., 1991. Urban GIS applications. In *Geographical Information Systems Principles and Applications,* edited by D. Maguire, M. Goodchild and D. Rhind (Essex, England: Longman Scientific & Technical), pp. 247-260.

Weinschenk, S., Jamar, P., and Yeo, S., 1997. GUI design essentials. *Wiley corp*., New York.

Worboys, M. F., 1995. GIS – A Computing Perspective. *Taylor & Francis*, London.