Dissemination of temporal geospatial data via a national Web atlas within in a Geospatial Data Infrastructure

Module 13 – Elective 3: Time in spatial data handling

October 1999

E. Mac Gillavry
GFM.2
Supervising staff: M.J. Kraak

[Geospatial] data sharing makes sense for the simple reason that there is only one Earth, and we share it.

Kurt Buehler and Lance McKee, Open GIS Consortium Technical Committee

Foreword

This report is submitted in partial fulfilment of the requirements for the assessment of Module 13 (Elective 3: Time in spatial data handling) of the curriculum of the Master of Science course in Geoinformatics at the International Institute for Aerospace Survey and Earth Sciences (ITC) in Enschede, the Netherlands.

Contents

Chapter 1

Introduction

One of the main assets of a society is its environment, both the natural environment and the physical environment. In order to make decisions about this environment, good geographical information is necessary. This geographic information is derived from geospatial data, data describing real-world phenomena directly or indirectly associated with a location relative to the surface of the Earth.

Geospatial data have been collected in digital form for more than 30 years. The overall rate of collection increases rapidly both with advances in technologies such as high-resolution satellite-borne imaging systems and global positioning systems (GPSs), and with the growing number of people and organisations who are collecting and using geospatial data. That number will continue to grow with the growing awareness among information technologists that indexing data by location is a fundamental way to organise and use digital data.

The integration of geospatial data is increasingly important because of growing environmental concerns, pressures on national governments and businesses to perform more efficiently, and simply because of the existence of a rapidly growing body of useful geospatial data and of geo-processing tools. A National Geospatial Data Infrastructure (NGDI) enables this integration of geospatial data from. The national geospatial data community, suppliers and users, the public and private sector and the academic community must agree on the nature and elements of a NGDI and how to develop it. However, there is no straightforward, universal, concise definition of a NGDI, for it may refer to different concepts. For the scope of this paper the following definition is pursued:

- *A National Geospatial Data Infrastructure (NGDI) is a set of institutional, technical and economical arrangements to enhance the availability of correct, up-to-date, to-the-point and integrated geospatial data related to the geography of one particular country, timely and at an affordable price to support decision making processes*.

A National Clearinghouse for Geospatial Data (NCGD) is the binding factor in this NGDI and provides an interface to this infrastructure. Via the NCGD clients are able to access, query and obtain the various geospatial data sources related to the geography of one particular country. For years, national atlases have provided this visual, structured, cohesive, and contemporary view of a country's geography. In this paper the concept of a national atlas is updated to fit into the Web environment. The Web has become – or in a not to distant future will become – the primary medium for disseminating information to a wide audience. This makes the Web most appropriate for a national atlas to reach a nation-wide audience. As the contents depend directly on geospatial data providers participating in the National Geospatial Data Infrastructure, a national Web atlas would provide a visual, structured, cohesive, and contemporary view of a country's geography to access, query, and obtain the various geospatial data sources related to the geography of one particular country.

This new concept of a national Web atlas as part of the NGDI is described in chapter two as well as the various options for the role this national Web atlas plays with regard to the NCGD. The third chapter focuses on the temporal aspects involved in such a setup, as in the definition of a NGDI *time* is frequently referred to. Problems of history-oriented databases are addressed as well as their possible solutions. Solutions are derived from various sources and discussed to define in which way they may contribute in solving the encountered problems. The final chapter provides a summary of the paper and gives conclusions that can be drawn to set-up a national Web atlas as part of a NGDI enabling the timely provision of good and up-to-date and temporal geographical information in order to make decisions about a society's natural and physical environment.

Chapter 2

NGDI: a visual approach

Introduction

The Web provides new challenges and opportunities for the representation, dissemination, and communication of information. Easy access to the Web provides an array of data, allowing companies to gather and disseminate dynamic internal and external information.

Because the Web is a heterogeneous environment of platforms and applications, interoperability at all levels has to be ensured, e.g. by means of standards for data exchange. A National Clearinghouse for Geospatial Data (NCGD) allows geospatial data providers to show what geospatial data exist related to a country's geography, the conditions of these data and instructions for accessing the data to exchange (Bishr & Radwan, 1998, p.14). To make clear to the user community the value of geospatial data sets available through the National Geospatial Data Infrastructure (NGDI) before the actual exchange takes place, descriptions of the datasets (metadata) should be provided. Each geospatial data provider describes the available data and provides these metadata to the NCGD. From the client's point of view the NCGD is the entry point of the NGDI when searching for geospatial data related to one country. The NCGD provides clients with a metadata database search engine to access, query, and obtain geospatial datasets from various sources.

Maps that visualise geospatial data can play a role at this point in the National Geospatial Data Infrastructure. On the Web maps are becoming more and more popular. Maps on the Web, the so-called Web maps can be classified in static and dynamic maps which are either view-only or interactive (URL 1).

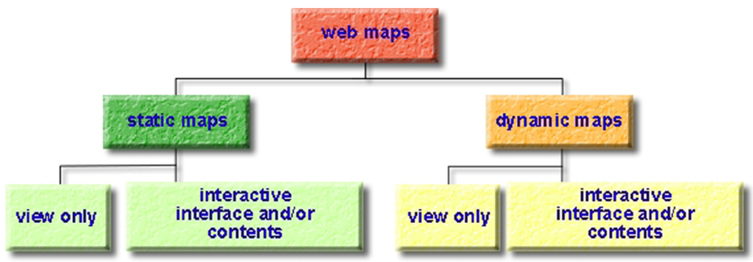

Figure 1. Classification of webmaps

Web maps can visualise geospatial data at a national scale: as a teaser to attract clients to download a specific data set, free of charge or at a certain price. Here, static, view-only Web maps play a role in advertising the different services on offer, both analogue and digital. Static, interactive maps can be provided to give an idea of the possibilities of digital vector data. Clients can switch data layers on and off. Instead of downloading a sample data set and viewing it with a GIS on the client-side, the Web site might provide a tool for server-side, online browsing, thus circumventing the necessity to download data.

Another role Web maps can play in the NGDI is providing a visual, interactive tool to select geospatial datasets, based on their spatial coverage. In the previous paragraph, Web maps did not come into play until the last stage of dataset selection. In this paragraph however, Web maps are at the forefront of this process. These Web maps provide interfaces and act as gateways between the Web and the various, distributed sources for geospatial data. This enables a client to use a Web browser to conduct powerful searches simultaneously at various geospatial data sources. The spatial search interface allows clients to define the spatial coverage of their query by selecting an area. Another way to use interactive Web maps is as a *clickable image*. Every hotspot in the image refers to one map sheet: the clickable image as an index sheet to the analogue products.

Furthermore, dynamic maps play a role. Dynamic, view only maps can be a means to show the possibilities of digital vector data layers, showing a consecutive layering of themes. Here, only a presentation of the possibilities of the layer structure is enabled, without any interaction, as is provided by the static, interactive map type. This is an option for that part of the Web site open to the general public, whereas the static, interactive map type might be a tool on the Web pages for the professional clientele. A dynamic map can also be a tool to simulate a fly-through to give insight in the 3D data on offer. Dynamic, interactive Web maps provide tools to explore geospatial datasets in a Web environment.

Exploiting the potential of Web maps to visualise the country's geography at the National Clearinghouse for Geospatial Data (NCGD) could result in a national Web atlas as part of the National Geospatial Data Infrastructure (NGDI). As a NCGD plays a central role within this NGDI by enabling clients to access, query and obtain the various distributed geospatial data sources, a national Web atlas – deriving its contents directly from participants in the NGDI – would provide a visual, structured, cohesive, and contemporary view of a country's geography.

National Web atlas

For years, national atlases have provided this visual, structured, cohesive, and contemporary view of a country's geography, whilst at the same time showing its state of the art technology, skills, and abilities in cartography. A national Web atlas might perform this same function in the information age. A national Web atlas would provide a cartographic framework to represent a country's geography, deriving its contents from the geospatial data providers participating in the National Geospatial Data Infrastructure.

This cartographic framework implies pre-designed views for particular maps. In practise this means that the page layout, such as location of the title and legend, as well as options for interaction are fixed, and only the visualisation of the data changes. For instance a geographic unit might fall in a different class or the size of a symbol is increased when the dataset changes over time. The options for interaction may be for the user to re-classify the dataset, change the representation or map-type (re-expression).

Most geospatial data providers sympathise with the idea of a national Web atlas: it is a perfect means for them to advertise their services on a national scale. However, they do not want put any effort in it, especially since it has become practice to keep the data at the data provider's Web site. National atlases, be they in paper or electronic form, can be major drains on resources, because of the investments in time and data involved as used to be common practice with a National Atlas Office. A drawback of this database-generated national Web atlas scenario would be that the contents are not comprehensive. The available geospatial datasets would not cover all the themes related to the country's geography, as the national Web atlas would only comprise the datasets offered by the participants of the NGDI.

National Web atlas: its position within the NGDI

The role of the national Web atlas towards the NGDI can have several forms. In one scenario, a national Web atlas is the visual interface of the National Clearinghouse for Geospatial Data to the NGDI. In this case the national Web atlas is an integral aspect of the NGDI and forms one, holistic system with the Clearinghouse. The national Web atlas provides the cartographic technology, skills, and abilities to visualise the various geospatial datasets on a national scale as provided by the participants. Furthermore, the national Web atlas provides tools to search both the national Web atlas itself and the Clearinghouse for geospatial data and facilitating e-commerce between the clients of the Clearinghouse and the data providers.

A more modest approach to the national Web atlas is the atlas as a means of teaser, an advertisement of the data that can be accessed through the NGDI. In this scenario the Web atlas provides only insight into the data at a national scale: the Web atlas as a collection of various thematic maps, just like the conventional national atlas. In this scenario, the NCGD and the Web atlas co-exist: when clients require other datasets, they need to find their way to the NCGD themselves. The national Web atlas differs from the conventional paper atlas in one important aspect: the currency of the data on which the thematic maps are based. As soon as the concept has been implemented, then the production time for new Web maps is far shorter than that of paper maps, let alone a whole atlas: the Web atlas is fully database-driven using direct links to the data providers.

A third approach combines these extremes. Here, the NCGD and the Web atlas are interchangeable interfaces depending on the client's objectives and the client profile with regard to geographical skills. Whether clients first approach the Web atlas or the Clearinghouse, they can visualise the small-scale datasets found in the Clearinghouse by entering the Web atlas and they can look for more detailed information in the Clearinghouse once they found interesting themes on a national scale in the Web atlas.

These three scenarios can co-exist fairly easily with the concept of the "Stufenatlanten". This concept was first applied to German school atlases. Depending on the pupil's grade, the school atlas provides more or less functionality and help options. Applying this concept to the national Web atlas would mean that depending on the client's profile, more or less functionalities are made available. In the next chapter the concept of a national Web atlas is taken further, with special attention to the temporal aspects.

Chapter 3

Tools to deal with time in Web atlases

Needs for time in atlases

Within the conceptual structure of the national Web atlas and NGDI *time* plays a significant role. There are three aspects to time in this structure. A national atlas initiative should be able to present its clients a picture of the *present* geography of a country, and maintain this adage into the *future*. With the scenario of a national Web atlas as part of a NGDI, clients should also be able to access outdated geospatial datasets e.g. for monitoring studies. Based on these outdated datasets the national atlas would be able to present its clients a picture of the *past* geography of a country.

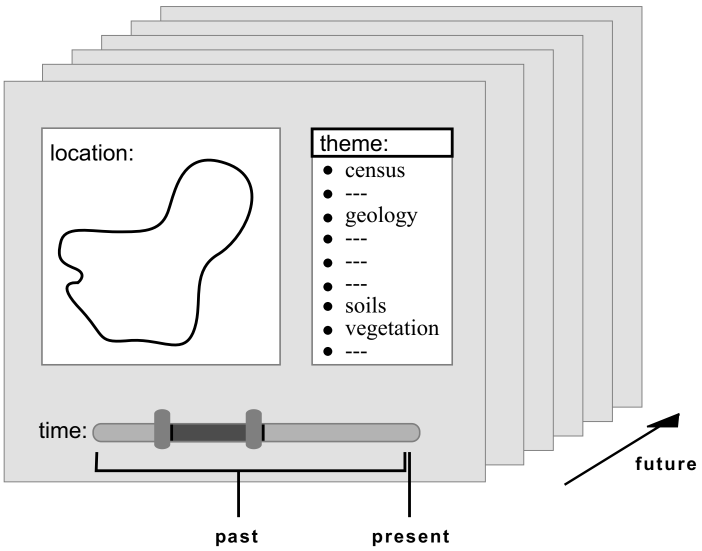

Figure 2. Time in Web atlases

Nowadays, national atlases in general are in danger of loosing their relevance if not kept up-to-date continuously. Clients would like at least to be able to obtain an impression of the *present* situation of their country. This temporal requirement either requires a permanent National Atlas Office or – as is apparent from the concept outlined before – a national Web atlas with input from permanent, direct, on-line links to the geospatial data suppliers participating in the NGDI who keep their own databases up-to-date.

Another manifestation of time with regard to currency is the fact that the Web is ever expanding over time. Every minute new information is made available to a global audience through this medium. So, it would be good practice if the national Web atlas not only kept track of updated contents of the datasets it already contains, but also discovers new datasets to keep up-to-date with the geospatial contents of the Web once the national Web atlas is implemented. In this way the national Web atlas truly provides an up-to-date view of the country not only at the time of its initial implementation, but also in the *future*.

As up-to-date datasets are available through the national Web atlas at present and new ones become available in the future the problem arises what to do with the outdated datasets. One option would be to keep them at the data provider. This option is only viable if the data provider itself has an advantage in storing these datasets, as for example a National Bureau of Statistics. This provider has an interest in outdated datasets, for they enable the Bureau to make trend analyses. However, for other participants in the NGDI, this maintenance of outdated datasets it beyond their interest or mandate. In this case, the solution could be to maintain the datasets at the server of the national Web atlas. Now that a distributed, history-oriented database is available, time plays a role as input when performing a temporal search for appropriate datasets. For example, the client of the NCGD would like to have the latest information about a particular theme as input for analysis, and the Web atlas user would like to know about the rate of migration in 1975. In this chapter tools are described that satisfy the temporal requirements of a national Web atlas.

Automatic updating

In discussions about the Web “easy-to-update” is always referred to when promoting its use. However, when it comes to Web sites with database-generated contents, especially when geospatial databases are concerned, one runs into problems. How to keep the database-generated site contents up-to-date? Translating this challenge the realm of geospatial data, the question should be rephrased into the question how to keep database-generated Web maps up-to-date? If the participants in the NGDI, both geospatial data providers and the organisation of the national Web atlas, do not have an organisational infrastructure in place to take care of up-dating the data, even the Web will not help.

In order to derive the latest geospatial data from the participants in the NGDI, the national Web atlas depends on the data descriptions, the metadata. To bypass the absence of efforts, time and money, the implementation of the national Web atlas should provide a mechanism to check the metadata of the geospatial databases residing with the providers. This implementation does not put any burden on the data providers. It assumes they are willing to provide access to a sub-set of their data as direct, online input for the national Web atlas.

Such a mechanism for checking Web site contents is a Web agent. These computer systems behave in such a way that, as soon as changes or updates are observed in the site contents at the server of the data provider, relevant data is uploaded onto the Web server of the national Web atlas and subsequently processed. Using this mechanism, clients are always assured of the currency of the metadata and maps.

Agent technology

An (intelligent) agent is a computer system situated in an environment and it is capable of *flexible autonomous* action in order to meet its design objectives (Jennings & Wooldridge, 1998, p.4). An agent is autonomous because it acts without direct intervention of humans (or other agents) and has control over its own actions and internal state. An agent is flexible, because it is responsive, proactive, and social. The agent perceives its environment and *responds* to changes that occur in it. Moreover, the agent exhibits opportunistic, goal-directed behaviour and takes initiative: it is *proactive*. The agent interacts at its own will with other agents and humans both to complete its own problem solving tasks and to help others with their tasks: it is *social*. Agents also exhibit some number of attributes, the key ones of which are *learning*, *co-operation*, and *mobility* (URL 2).

Intelligent agents are a fundamental enabling technology for brokerage (Foss, 1998, p.107). Therefore the agent paradigm is very suitable to adapt to the clearinghouse concept. They have the characteristic needed to support this dynamic service environment: mobility between a user and other users and across platforms; autonomy to initiate and perform tasks; representation of their owners’ interest across a network; flexibility and adaptiveness, customising their actions to the evolving environment and communicativeness.

Different agents perform different tasks within the clearinghouse environment. With regard to the problem of updating, Web agents are important. A Web agent is a special class of intelligent agents, a computer system designed to manage, manipulate, or collate information from many distributed sources (Nwana & Ndumu, 1998, p.38). One type of Web agents includes specific site watchers that notify its owner of site updates. Typical static Web agents are embedded within a Web browser and use a host of Internet management tools such as search engines to gather the information. Since the key problem with these static Web agents is to keep their indexes up-to-date in an environment prone to change, future Web agents will be mostly mobile: they navigate the Web and store its topology in a database to be queried by the user.

Site watchers, a quick overview

There are some commercial Web agents. InformationMiner4U (URL 3) is designed to conduct an automatic watch on the Web in order to detect changes, appearances, or disappearances of information. It is a server watch system for Internet/Intranet that discovers and watches the evolution of what has been discovered continuously. Every time something happens, according to predefined parameters, it sends an email with summarised information, indicating the nature of movements of information (apparition, disappearance, and semantic modification).

javElink (URL 4) is a complete Web page change monitoring service. By simply listing the Web pages to be monitored, e.g. a bookmark file), and javElink reviews every page, every day. The private checklist can be logged on whenever appropriate. It discovers new, different, or deleted text, then scores each page by the degree of change. It shows a "thumbnail" view of any page with recent changes highlighted. The e-mail notification may include all or only part of the changed text, and deleted text according to user-defined configurations. Some sites may change by the minute (such as weather monitoring sites). These sites should be visited directly: javElink only alerts the user to changes since the previous day.

Katipo (URL 5) finds any documents that have changed since the last time it was visited by the user. It writes a report file listing these documents. Katipo uses a special feature of the Web that allows it to check for changes without loading the whole document. This makes the agent really efficient. Katipo does not download any documents; it just checks their characteristics.

Morning Paper (URL 6) automatically visits user-defined Web sites every so often to find out what's new, and presents a summary of what's new on each page as part of a "newspaper" displayed in a Web browser. Morning paper keeps the client up-to-date on what's new and interesting, and provides convenient links to visit the pages that have changed.

NetMind (URL 7) has developed a complete product line to service a variety of needs to notify people of updates:

- Mind-it: to track business and personal information on the Internet
- Webmaster Program: to integrate NetMind's buttons into the user's Web sites, allowing people to easily track changes to any page within that Web site.
- Enterprise Minder: server product that works behind a firewall and allows employees to easily track changes to any page of a corporate intranet or the Internet.
- Minder for Partners: to couple Enterprise Minder with customisation and service, so that E-commerce and E-publishing Web sites can integrate personal, customised tracking solutions.

This overview of commercially available site watchers shows some overall functionalities of this type of Web agents:

- Observation: the Web agent watches user-defined or explored URLs for change
- Exploration: the Web agent explores the Web looking for new sites
- Personalisation: the Web agent explores the Web based on the personal profile of the client
- Notification: the Web agent tells the client that something happened: "The web page at http://www.ingetech.com has changed!"
- Details: the Web agent tells all of what changed: "All changes found today at http://www.ingetech.com..."
- Extraction: the Web agent tells what is relevant to the client: "Changes at http://www.ingetech.com meeting criteria…"
- Visualisation: the Web agent highlights the changes or displays the latest relevant information
- Archival: the Web agent documents all changes "The change history of http://www.ingetech.com is…"

In the next paragraph these overall characteristics of site watchers are discussed in the context of a National Geospatial Data Infrastructure, where they provide a mechanism for updating and expanding the national Web atlas.

Site watchers, their relevance to the NGDI

Most site watchers are designed to communicate changes in (database-generated) site contents directly to the end-user. This could also be relevant to the implementation of a national Web atlas, though here the changes should be communicated to the information broker, the national Web atlas, to update the maps based on the new database values. The NetMind (URL 7) software might facilitate this mechanism, for it has a more server-oriented set-up, whereas most site watchers discussed are client-oriented.

In order to evaluate the relevance of the other site watchers, one should see the server of the national Web atlas, in this implementation, as a client to the servers of the geospatial data providers. Site watchers observe these servers for changes in the metadata to ensure the currency of the national Web atlas. Furthermore, they explore the Web for new sites. These tasks are performed based on the requirements set by the participants of the NGDI, making up the user profile of the national Web atlas as the client. This ensures future currency after implementation. Web agents can notify the national Web atlas of changes, the details of all these changes and extract only the relevant changes. Site watchers use e-mail notification for this purpose, as do InformationMiner4U (URL3) and javElink (URL 4). Some site watchers visualise, highlight these changes or display the latest relevant information. The Morning Paper (URL 6) service is even more interesting, for it appeals to the idea of targeting a large audience and it notifies its owner using Web pages instead of e-mail. The visualisation functionality has to be extended to visualise not the changes themselves (though this maybe another, advanced functionality), but the up-to-date values of the geospatial database so the map is refreshed. The Katipo (URL 5) site watcher is also worth paying attention. Katipo checks Web sites for changes without loading the whole document, it just checks their characteristics, their metadata. The last functionality described in the previous paragraph is archival. Site watchers document the changes in site contents over time, thus describing the history of a Web site. This appeals to the necessity for a national Web atlas to store the past.

This discussion shows that Web agents, especially site watchers, are indeed very useful in the context of the National Geospatial Data Infrastructure, for they meet the requirements for time as set before in the introductory paragraph of this chapter. The usefulness of Web agents in this context is exemplified in the next paragraph.

Geospatial Metadata Server: a geospatial example

One application of Web agents in a NGDI is the Geospatial Metadata Server (GMS). This server-side, on-line software system merges existing technologies such as a HTTP server, a relational metadata database and a GIS into a unified system that facilitates the on-line access and conversion of geospatial data on the Web (Pradhan & Gittings, 1998, p.39; URL 8). Metadata text files are uploaded and subsequently parsed for entry into a database. The entry, retrieval, updating, and removal of metadata is performed on-line. Querying of the database can be defined by combining parameters for keyword, date and co-ordinate. Datasets resulting from a query can be converted into user-defined formats by means of data conversion utilities integrated into the database. A Web agent, GEOBOT, integrated into the database, explores an entire Web site looking for geospatial data. Arc/Info coverages can be viewed on-line.

The updating of the metadata in this system has to be performed by the geospatial data provider (Pradhan & Gittings, 1998, p.42). This process is facilitated considerably using HTML forms or a Java programmed interface. This update is directly available for clients to query. However, one of the requirements set in this paper, is that the provider should not have to put any effort in the maintenance of the metadata database at the NCGD/national Web atlas. Therefore, the updating mechanism designed for the GMS does not fully comply with the requirement set here.

One of the metadata database entries for each geospatial dataset contains the URL of the Web site where it resides. It could happen that the URL is not referring to the geospatial dataset itself, but to a Web site containing lots of datasets. The GEOBOT Web agent scans the whole Web site page by page for links to geospatial datasets by looking at file extensions (Pradhan & Gittings, 1998, p.50). These links are then entered in the metadata database and can be queried by the GMS. Here, it shows that Web agents perform an important role in the context of a NGDI.

Web agents do not provide a mechanism to comply with the requirement to access outdated geospatial datasets. However, this is very important as is apparent from this example of the Geospatial Metadata Server as well. In the next paragraph a tool is described to cater for this.

Temporal queries in a NGDI

Introduction

Clients to a national Web atlas are not only interested in a contemporary view of their country's geography. There is a growing interest for the historical aspects of geography. This is reflected in the geosciences. Geospatial databases have existed for a long time, though not until recently attention in research is given to their historical dimension (Peuquet, 1997, p.91). Three trends have driven this attention. First, there is a growing research interest in the analysis of geospatial patterns over time. Second, more and more data is collected of the same area and the same attributes over time (Jensen, 1996). Third, computing power and memory capacity increase every time, while the cost of memory decreases.

To perform temporal geospatial analysis, clients should have access to temporal datasets. A search engine should therefore have a functionality to perform temporal querying of the metadata database. The GMS descibed in the previous paragraph provides this functionality (URL 8). Clients select the time period for which they require datasets. The actual date entry is achieved with the use of drop-down lists.

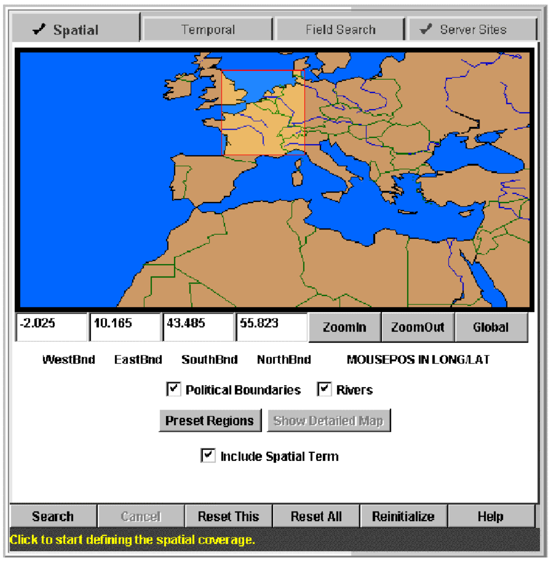

Figure 3. Interface to the Geospatial Data Clearinghouse

An advanced visual interface to a geospatial metadata search engine is provided at the Web site of the Geospatial Data Clearinghouse of the FGDC. This Web site (URL 9) is a collection of over 100 data servers. These collections can be searched based on their metadata through a single interface.

The Java-based interface uses an applet allowing the client to construct and submit a query based on the spatial and temporal coverages of the data, keywords in fields of the metadata and the server sites that are to be searched. The tabs allow clients to switch between panels allowing them to specify different portions of the query. The panel "✓ Temporal" allows them to specify a single date or a range of dates with which to search the metadata records.

Temporal queries, a typification

In order for clients to search geospatial datasets based on temporal metadata, the database has to be constructed to facilitate storing of temporal data. The first step in database design is the construction of a reference model or external model. The user view on subsets of the real world information has to be defined and described. In this context, this is accomplished by recording the questions related to time to be asked of the system in order to retrieve appropriate datasets:

- Find a dataset of the Netherlands about the population size of 1998
- Find all the datasets of the Netherlands about the population size from 1995 to 1997

These questions are input into the query interface of either the national Web atlas or the NCGD in order to retrieve appropriate datasets. However, these questions may result based on temporal geospatial queries put to the national Web atlas:

- Find the datasets that show the same annual increase in migration as in 1975
- Find the datasets showing the rate of unemployment in the five years following the years that have an annual increase of migration greater than 105%

Since these latter questions are more exploratory in nature, they will not be addressed until the next module. This reference model focuses only on questions related to dataset retrieval based on querying a metadata database with straightforward, temporal parameters.

Constructing a database for temporal querying

With hindsight to the logical model, geospatial entities are represented in a relational database using the topological vector approach. In the first place, the majority of work in temporal databases is based on the relational model (Snodgrass, 1992, p.34). While time passes, the two other attributes of an entity stored in a database may change. The geometry of an entity may change through time. At a national scale one can think of the annexation of one municipality by another or the movement of sandbanks along the coast. In the amendment vector approach the time when the change occurred is recorded as an attribute of each vector (Peuquet, 1997, p.95). This time attribute is referred to as *transaction time*. The transaction time of an entity is the time when the fact is current in the database and may be retrieved (Jensen *et al*., 1994, p.3). Transaction times are system-generated. The recording of the transaction time allows the integrity of individual entities, their components, and the topology to be explicitly maintained over time.

Also the thematic attribute of an entity changes over time as e.g. the population density per municipality. To track the history of thematic attributes, each entity is represented by a unique identifier and a sequence of attribute values (Peuquet, 1997, p.95). This time attribute is referred to as *valid time*. The valid time is the time when a fact is true in the modelled reality (Jensen *et al*., 1994, p.3). A fact may have any number of instances and time intervals. Valid times are user-defined.

If the geometric and thematic attributes are changing at different times and at different rates, maintaining the identity of individual entities becomes difficult (Peuquet, 1997, p.95). Both temporal and geometric attributes are required to identify an individual entity. The temporal attributes are utilised twice: first to identify an entity and then to select a particular thematic value associated with that entity.

A data model supporting both valid time and transaction time is termed a *bitemporal* data model. The design of a metadata database should reflect the contents of such a data model. Especially the valid time attribute is important in temporal querying. It seems as if transaction time would be an attribute useful for automatic updating. Every time there is a new transaction time, the latest data is entered. However, this does not need to be the case. Indeed, new data is entered, but not necessarily the latest data: *retrospective updating* (Peuquet, 1997, p.101; Snodgrass, 1992, p.25). Therefore, the agent-oriented approach towards updating does not lose its validity. In the next few paragraphs efforts are discussed that address the temporal aspects of metadata database design, with a special focus on the *valid time* attribute, which is most important for temporal queries. First, attention is directed towards the CSDGM standard applied in the example of the Geospatial Metadata Server discussed before.

The Content Standard for Digital Geospatial Metadata

The Content Standard for Digital Geospatial Metadata was developed by the US FGDC to identify and define the metadata elements used to provide a common set of terminology and definitions for the documentation of digital geospatial data sets for many purposes. An overview of the important aspects of this standard is available in Appendix A.

Time Period Information is provided in section 9 of the CSDGM. This section states the temporal information: the date and time of an event. This information plays a role in determining the fitness for use of a geospatial dataset and is used by other sections, e.g. Section 5. This section, “Entity and Attribute Information”, provides details about the information content of the data set, including the entity types, their attributes, and the domains from which attribute values may be assigned. Subsection 5.1.2.5, Beginning Date of Attribute Values, stores the earliest or only date for which the attribute values is current. In cases when a range of dates are provided, this is the earliest date for which the information is valid. Subsection 5.1.2.6, Ending Date of Attribute Values, stores the latest date for which the information is current. This subsection is used in cases when a range of dates is provided. Here the *valid time* is recorded in the metadata.

ENV 12657 Geographic Information – Metadata

Geographic Information documents prepared by Technical Committee CEN/TC 287 were adopted as European Pre-standards (ENVs). ENVs (Euro-Norme Voluntaire) have been established as prospective standards for provisional application in technical fields where the innovation rate is high or where there is an urgent need for guidance and primarily where the safety of persons or goods is not involved.

The work programme of TC 287 is divided into four main parts: fundamentals, data description, referencing and processing. Under the main heading of referencing, one of the three work items is Work Item 287012: Referencing – Time. This work item has now been suspended in CEN/TC 287. The result of the work of ISO/TC 211 on this topic will be considered for implementation as an EN.

```
ENTITY temporal_extent
SUBTYPE OF (extent);
temporal_extent_description    : OPTIONAL STRING;
has_period_range_details       : OPTIONAL period_range_details;
WHERE
WR1: EXISTS (temporal_extent_description) OR EXISTS
(has_period_range_details);
END_ENTITY;
ENTITY period_range_details;
period_start_date                   : date;
period_end_date                  : date;
      period_range_date_qualifier    : OPTIONAL STRING;
(* WHERE
WR1: period_start_date <= period_end_date; *)
END_ENTITY;
```

Figure 4. ENV 12657: valid time

ENV 12657 (URL 11) defines a conceptual schema for metadata for geographic datasets. Metadata is data about datasets. It includes information about the content, representation, extent (both geometric and temporal), spatial reference system, quality, and administration of the dataset. ENV 12657 identifies those items that are mandatory for describing geographic datasets – the minimum set of metadata. One of these items, or entities, is the *temporal extent*. Another one is the *period range details*. So also this metadata standard supports the storage of *valid time*.

The Open GIS Consortium

The Open GIS Consortium (OGC) is a not for profit membership organisation, whose mission is the full integration of geospatial data and geo-processing resources into mainstream computing and the widespread use of interoperable, commercial geo-processing software throughout the global information infrastructure. To achieve this vision, OGC works to involve the world's developers and users of geographic information resources - including vendors, integrators, academia, government agencies, and standards organisations - in collaborative development of interoperable geo-processing technology specifications, and to promote the delivery and use of interoperable geo-processing products based on these specifications.

With regard to standardisation initiatives, OGC has started co-operating with Technical Committee 211 of the International Standardisation Organisation, ISO/TC211 (URL12). Until then, OGC itself had not developed a temporal schema (ISO/TC211 work item CD 10/97). OGC has decided to “outsource” the part of their work programs related to this temporal sub-schema to the ISO/TC211 organisation, for resources available in each organisation are complementary in some areas and could be leveraged for mutual benefit.

Wrap-up

There are various initiatives around the globe standardising metadata descriptions, each encompassing another level of scale. The FGDC standard is only enforced within the USA. The ENV standards are followed only within the European Union. The OGC tries to accomplish a global initiative together with Technical Committee 211 of the ISO. All of these standardisation initiatives have incorporated temporal attributes in the metadata descriptions. As the historical dimension of geospatial datasets becomes important in the implementation of a national Web atlas, there seems to be enough ground to work upon.

Chapter
4

Summary and conclusions

This paper addressed the setup of a national Web atlas together with a National Clearinghouse for Geospatial Data (NCGD). In this scenario the contents of a national Web atlas depend on geospatial data providers, participating in the National Geospatial Data Infrastructure (NGDI). The second chapter described this setup as well as various options for the role a national Web atlas may play within a NGDI with regard to the NCGD. Three options were described for this setup:

- Holistic system
- Co-existent systems
- Interchangeable system

The third chapter focused on the temporal aspects involved in such a setup. How to keep the contents of the national Web atlas up-to-date at the time of implementation and how to maintain and expand the contents this in the future and the past and how to deal with outdated datasets? These problems were addressed as well as possible solutions. These solutions were derived from various sources and discussed to define in which way they may contribute in solving the encountered problems. Agent technology provides a promising solution for the currency problem at present and to maintain and expand the contents in the future. Site watchers, a special type of Web agents, have functionalities to watch user-defined or explored URLs for change, to explore itself the Web looking for Web sites of interest to the client based on the personal profile of the client. The Web agent tells clients that something happened, tells all of what changed, tells what is relevant to the client, highlights the changes or displays the latest relevant information, and documents all changes. In this context the services provided by Morning Paper and NetMind are most interesting. The Morning Paper service appeals to the idea of targeting a large audience. The NetMind software might facilitate the provision of the update service from the geospatial data providers to the server of the national Web atlas, which is in this configuration the client.. The visualisation functionality has to be extended in the context of the national Web atlas in order to visualise not the changes themselves (though this maybe another, advanced functionality), but the up-to-date database values, so the database-generated map is refreshed. The application of Web agents was exemplified in a discussion of the Geospatial Metadata Server.

From this example it became once more apparent that there is a need to access outdated datasets. In order to facilitate temporal querying of geospatial datasets in a NCGD, it is necessary to have the temporal aspects of the datasets stored in a metadata database. For temporal queries the *valid time* attribute is most important. Various initiatives around the globe each encompassing another level of scale were discussed. The FGDC standard is only enforced within the USA. The ENV standards are followed only within the European Union. The OGC tries to accomplish a global initiative. All of these standardisation initiatives have incorporated temporal attributes in the metadata descriptions. As the historical dimension of geospatial datasets becomes important in the implementation of a national Web atlas, there seems to be enough ground to work upon.

Web agents and metadata standards encompassing temporal aatributes provide useful mechanisms to deal with the several aspects of time in the context of a national Web atlas.

Appendix A : CSDGM[1]

Organisation of the Standard

Calendar Dates (Years, Months, and Days)

A.D. Era to December 31, 9999 A.D. – Values for day and month of year, and for years, shall follow the calendar date convention (general forms of YYYY for years; YYYYMM for month of a year (with month being expressed as an integer), and YYYYMMDD for a day of the year) specified in American National Standards Institute, 1986, Representation for calendar date and ordinal date for information interchange (ANSI X3.30-1985): New York, American National Standards Institute (adopted as Federal Information Processing Standard 4-1).

- B.C. Era to 9999 B.C.: Values for day and month of year, and for years, shall follow the calendar date convention, preceded by the lower case letters "bc" (general forms of bcYYYY for years; bcYYYYMM for month of a year (with month being expressed as an integer), and bcYYYYMMDD for a day of the year).

- B.C. Era before 9999 B.C.: Values for the year shall consist of as many numeric characters as needed to represent the number of the year B.C., preceded by lower case letters “cc” (general form of ccYYYYYYY...).

- A.D. Era after 9999 A.D.: Values for the year shall consist of as many numeric characters as needed to represent number of the year A.D., preceded by the lower case letters “cd” (general form of cdYYYYYYY...).

Time of Day (Hours, Minutes, and Seconds)

Because some geospatial data and related applications are sensitive to time of day information, three conventions are permitted. Only one convention shall be used for metadata for a data set. The conventions are:

- Local Time. For producers who wish to record time in local time, values shall follow the 24- hour timekeeping system for local time of day in the hours, minutes, seconds, and decimal fractions of a second (to the precision desired) without separators convention (general form of HHMMSSSS) specified in American National Standards Institute, 1986, Representations of local time of day for information interchange (ANSI X3.43-1986): New York, American National Standards Institute.

- Local Time with Time Differential Factor. For producers who wish to record time in local time and the relationship to Universal Time (Greenwich Mean Time), values shall follow the 24-hour timekeeping system for local time of day in hours, minutes, seconds, and decimal fractions of a second (to the resolution desired) without separators convention. This value shall be followed, without separators, by the time differential factor. The time differential factor expresses the difference in hours and minutes between local time and Universal Time. It is represented by a four-digit number preceded by a plus sign (+) or minus sign (-), indicating hours and minutes local time is ahead of or behind Universal Time, respectively. The general form is HHMMSSSSshhmm, where HHMMSSSS is the local time using 24-hour timekeeping (expressed to the precision desired), 's' is the plus or minus sign for the time differential factor, and hhmm is the time differential factor. (This option allows producers to record local time and time zone information. For example, Eastern Standard Time has a time differential factor of -0500, Central Standard Time has a time differential factor of -0600, Eastern Daylight Time has a time differential factor of -0400, and Central Daylight Time has a time differential factor of -0500.) This option is specified in American National Standards Institute, 1975, Representations of universal time, local time

[1] derived from (URL 10)

differentials, and United States time zone reference for information interchange (ANSI X3.51-1975): New York, American National Standards Institute.

- Universal Time (Greenwich Mean Time). For producers who wish to record time in Universal Time (Greenwich Mean Time), values shall follow the 24-hour timekeeping system for Universal Time of day in hours, minutes, seconds, and decimal fractions of a second (expressed to the precision desired) without separators convention, with the upper case letter “Z” directly following the low-order (or extreme right hand) time element of the 24-hour clock time expression. The general form is HHMMSSSSZ, where HHMMSSSS is Universal Time using 24-hour timekeeping, and Z is the letter “Z”. This option is specified in American National Standards Institute, 1975, Representations of universal time, local time differentials, and United States time zone reference for information interchange (ANSI X3.51-1975): New York, American National Standards Institute.

Entity and Attribute Information

Section 5 of the CSDGM, Entity and Attribute Information, provides details about the information content of the data set, including the entity types, their attributes, and the domains from which attribute values may be assigned.

Type: compound
Short Name: eainfo

5.1.2.5 Beginning Date of Attribute Values -- earliest or only date for which the attribute values are current. In cases when a range of dates are provided, this is the earliest date for which the information is valid.
Type: date
Domain: free date
Short Name: begdatea

5.1.2.6 Ending Date of Attribute Values -- latest date for which the information is current. Used in cases when a range of dates are provided.
Type: date
Domain: free date
Short Name: enddatea

Time Period Information

Section 9, Time Period Information, provides information about the date and time of an event. This section provides a means of stating temporal information, and is used by other sections of the metadata standard. This section is never used alone.

Type: compound
Short Name: timeinfo

Time_Period_Information =
[Single_Date/Time |
Multiple_Dates/Times |
Range_of_Dates/Times]

Single_Date/Time =
Calendar_Date +
(Time_of_Day)

Multiple_Dates/Times =
2{Single_Date/Time}n

Range_of_Dates/Times =
Beginning_Date +

(Beginning_Time) +
Ending_Date +
(Ending_Time)

9.1 Single Date/Time -- means of encoding a single date and time.
Type: compound
Short Name: sngdate

9.1.1 Calendar Date -- the year (and optionally month, or month and day).
Type: date
Domain: “Unknown” free date
Short Name: caldate

9.1.2 Time of Day -- the hour (and optionally minute, or minute and second) of the day.
Type: time
Domain: “Unknown” free time
Short Name: time

9.2 Multiple Dates/Times -- means of encoding multiple individual dates and times.
Type: compound
Short Name: mdattim

9.3 Range of Dates/Times -- means of encoding a range of dates and times.
Type: compound
Short Name: rngdates

9.3.1 Beginning Date -- the first year (and optionally month, or month and day) of the event.
Type: date
Domain: “Unknown” free date
Short Name: begdate

9.3.2 Beginning Time -- the first hour (and optionally minute, or minute and second) of the day for the event.
Type: time
Domain: “Unknown” free time
Short Name: begtime

9.3.3 Ending Date -- the last year (and optionally month, or month and day) for the event.
Type: date
Domain: “Unknown” “Present” free date
Short Name: enddate

9.3.4 Ending Time -- the last hour (and optionally minute, or minute and second) of the day for the event.
Type: time
Domain: “Unknown” free time
Short Name: endtime

Literature

[Foss, 1998]
Foss, J.D. (1998) Brokering the Info-Underworld. In: Jennings, N.R. & M.J. Wooldridge (eds.) Agent Technology. Foundations, Applications, and Markets. Berlin: Springer –Verlag. , p.105-124.

[Jennings & Wooldridge, 1998]
Jennings, N.R. & M. Wooldridge (1998) Applications of Intelligent Agents. In: Jennings, N.R. & M.J. Wooldridge (eds.) Agent Technology. Foundations, Applications, and Markets. Berlin: Springer –Verlag. , p.3-28.

[Jensen, 1996]
Jensen, J.R. (1996) Digital change detection. In: Introductory Digital Image Processing – A remote Sensing Perspective. Upper Saddle River: Prentice-Hall, 2nd edition, p.257-279.

[Jensen *et al*., 1994]
Jensen, Ch. S. *et al*. (1994) A consensus glossary of temporal database concepts. ARPA/NSF International Workshop on an Infrastructure for Temporal Databases. Earlier version published in SIGMOD Record 23 (1), p.52-64.

[Nwana & Ndumu, 1998]
Nwana, H.S. & D.T. Ndumu (1998) A Brief Introduction to Software Agent Technology. In: Jennings, N.R. & M.J. Wooldridge (eds.) Agent Technology. Foundations, Applications, and Markets. Berlin: Springer –Verlag. , p.29-48.

[Peuquet, 1997]
Peuquet, D.J. (1997) Time in GIS and geographical databases. In: Longley, P., M. Goodchild, D. Maguire & D. Rhind (editors) Geographical information systems: principles, techniques, management and applications. New York: J. Wiley & Sons, p.91-103.

[Pradhan & Gittings, 1998]
Pradhan, A. & B.M. Gitting (1998) Web-based geospatial data conversion tools. In: Gittings, B.M. (ed.) Innovations in GIS 6: Integrating Information Infrastructures with GI Technology. London: Taylor & Francis Ltd, p.39-52.

[Snodgrass, 1992]
Snodgrass, R.T. (1992) Temporal Databases. In: Frank, A.U., I. Campari & U. Formentini (eds.) Theory and Methods of Spatio-Temporal Reasoning in Geographic space. Pisa: Springer-Verlag, p.22-64.

Web links

[URL 1] **Web map classification**
http://www.itc.nl/~kraak/book/webcar1.htm

[URL 2] **Intro to intelligent agents**
http://www.cs.tcd.ie/research-grouips/aig/iag/pubreview.zip

[URL 3] **Information4U**
http://www.arisem.com/

[URL 4] **javElink**
http://www.javelink.com/

[URL 5] **Katipo**
http://www.vuw.ac.nz/~newbery/Katipo.html

[URL 6] **Morning Paper**
http://www.boutell.com/morning/

[URL 7] **NetMind**
http://www.netmind.com/

[URL 8] **Geospatial Metadata Server**
http://www.geo.ed.ac.uk/~anp/gms/main.htm

[URL 9] **Geospatial Data Clearinghouse**
http://clearinghouse1.fgdc.gov/FGDCgateway.html

[URL 10] **CSDGM standard**
http://www.fgdc.gov/metadata/csdgm/

[URL 11] **ENV 12657**
http://forum.afnor.fr/afnor/WORK/AFNOR/GPN2/Z13C/PUBLIC/DOC/metadata.txt

[URL 12] **OGC standardisation initiatives**
http://www.opengis.org/techno/standards.htm