Geodata – are they Accessible and Useful?

Marjan Čeh[1], Domen Smole[2], Tomaž Podobnikar[3]
[1] Faculty of Civil and Geodetic Engineering, University of Ljubljana
Ljubljana, Slovenia, mceh@fgg.uni-lj.si
[2] DFG CONSULTING, d.o.o., young researcher at Scientific Research Centre of the Slovenian Academy of Sciences and Arts
Ljubljana, Slovenia, domen.dfg@moj.net
[3] Institute of Anthropological and Spatial Studies at Scientific Research Centre of the Slovenian Academy of Sciences and Arts, Ljubljana, Slovenia
tomaz@zrc-sazu.si

SUMMARY

Experiences about solving a spatially-related problem have shown that best solutions and conclusions can be drawn when all kinds of existing spatial data are available. There are situations when data acquire has to be carried out. However, the questions that raises here are: whether the existent geodata is accessible, and whether it represents useful information to the user? This paper outlines some propositions on how to gain wider acceptance and use of geodata among professional and non-professional users. According to it at least two aspects play an important role when considering spatial data popularisation. Firstly, an infrastructure and technology that enables user-friendly access to the geodata has to be provided. In this respect, web technology and solutions based on semantic approaches and ontologies play an important role. Secondly, geodata should be useful to the users to such extent that the users return and take advantage of the geodata and geoservices again and again. In this regard, high quality and up-to-date geodata that would reflect current situation, acquired by utilizing modern and economic spatial data acquiring techniques, should be provided on regular basis.

KEYWORDS: *distributed computing environments, spatial data acquisition methods, user-oriented GIS, semantic GIS, Universal ontology of geographic space, semantic-geographic search*

INTRODUCTION

More than 80% of all information relates to some location on the Earth (Burrough & McDonnell, 1998) and only 10% of acquired and stored data is ever effectively exploited. On the other hand, we have witnessed rapid growth and popularisation of different types of communication media, especially the Internet and mobile phone networking. By utilizing modern means of communication, vast amounts of (spatially located) data could be accessed and used practically by anybody in a fast and economic way.

Internet is the biggest distributed system in the world and the World Wide Web is its most frequently used service. This is one of the reasons why the term “distributed GIS” is so often mistaken for the term “Web GIS”. However, today there are many commercially Web GIS solutions available. They may be distinguished according to the several factors. All of them should be considered with a great care when deciding which of many available GIS solutions to choose or buy. Beside certain technical and system oriented characteristics (for example, in which operating systems and Web browsers product operates, the number of spatial data formats the product supports, software and hardware deployment costs etc.), the need to be easy-to-use is very high on the list of user priorities (Limp, 2001).

For now, sophisticated off-the-shelf Web GIS solutions are expensive and require a substantial amount of knowledge for successful installation, configuration, fine-tuning etc. The migration of the specific GIS functionality to the core of a database management systems or the possibility of extending a database management system's business logic with spatial functionality in a combination with freely available standards technologies such as GML (Geography Markup Language), SVG (Scalable Vector Graphics), XSL (eXtensible Stylesheet Language) and some others, offer a competitive alternative (Plewe, 2002).

GEODATA ACCESSIBILITY

GIS software solutions are made to be used by the users. Providers often ask themselves how to make existent customers aware of having or using contemporary products and services they offer and how to attract new customers. The answer may be complex, but generally speaking, users must either gain a profit in using GIS technology or they must "simplify" some activities they do in everyday life (besides that it must be fun using GIS).

GIS technology has been taken online. However, experience has shown that not all aspects and benefits provided by the distributed systems have been fully addressed yet. We shall bear in mind that the use of spatially enabled data is still limited to the more or less skilled personnel and/or advanced users that understand GIS approaches and technology.

Application service providers

Application service providers (ASPs), companies mastering modern Web-oriented technologies, are expected to change the current situation. Spatially-oriented ASPs provide users with necessary hardware, software, spatial data and knowledge accessible over the Internet (Quartararo, 2002). One of the goals of ASPs is to prepare and present spatial information in an understandable way to everyone, especially to non-GIS experts. In return for monthly or per transaction based fee, users are given a chance to express their requirements and demands regarding spatial data/services in non technical terms, in terms they are familiar with. At the same time users are encouraged to hire GIS software owned by ASPs for the purpose of viewing the results of their requests and possibly deriving some additional information out of it. ASPs are aware of the fact that more the clients know about the GIS topics, the more demanding they are and the gain for providers rises. This is the reason why ASPs think that consulting, provision of education and general elucidation of GIS is so important.

An interesting possibility offered by new technologies is to allow users to maintain their own GIS projects on the Web. Not only viewing but adding and editing GIS oriented sites attract (more demanding) clients. The possibility of combining the data acquired on their own (for example with the help of a handheld GPS receiver, a digital camera, etc.) with the spatial data found on the Web may produce interesting results. Similarly, an idea of concurrent construction of one GIS "document" by several spatially dislocated authors may seem interesting enough and feasible too.

To ensure the previously mentioned combination of user's own data and spatially enabled data that can be found on the Web, *effective and simplified spatial data searching* must be provided. It means that looking for an available spatial data, for example on user's neighbourhood entities, should not explicitly require choosing the exact (usually coded) name representing the spatial database, which (according to the user's opinion) possibly contains information she/he is looking for. The development of *a user friendly spatial data search engine* would certainly make spatially aware datasets more valuable. For solving the problems of semantic interoperability of geodata there is crucial need to "invent" a universal semantic reference system. Some general clues have been given by Kuhn and Raubal (2003). An original attempt to this issue has been given also by Čeh (2003). The idea is semantic integration system which distincts between two independent and mutually connectable ontologies, namely semantic reference system of geographic space (developed Universal Ontology of Geographic Space – UOGS) and Ontology of GeoInformation Technology (OGIT) (Figure 1).

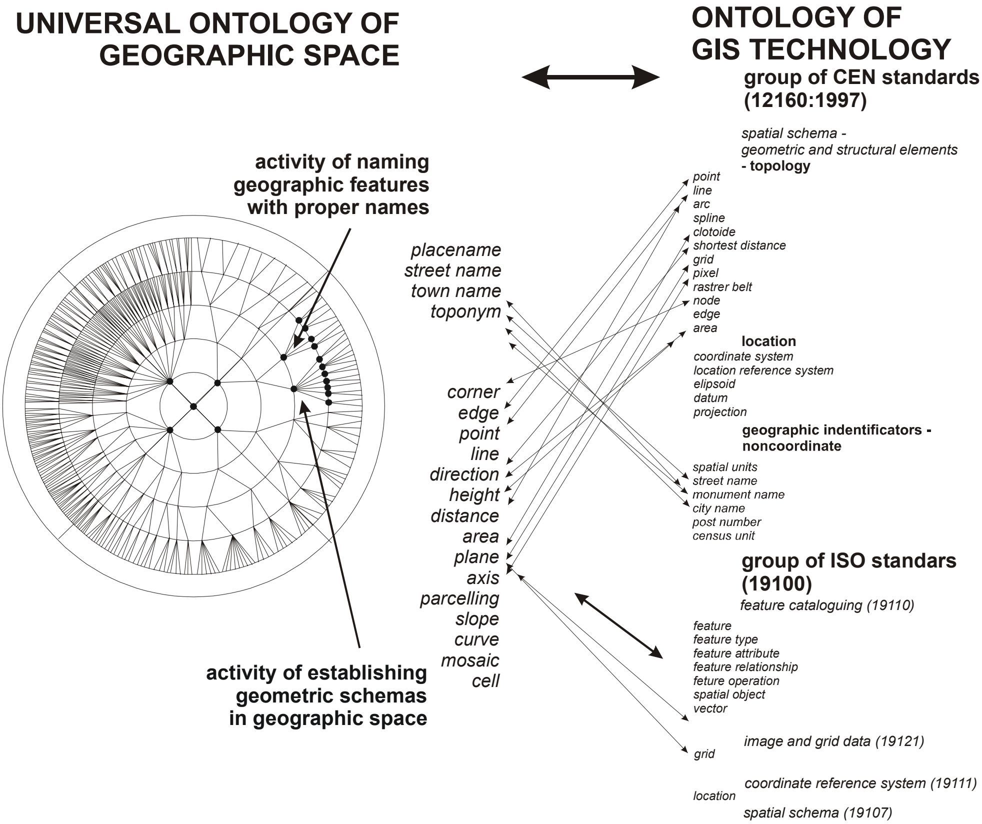

Figure 1: Development of effective interoperability of GIS and spatial databases requires holistic approach, which takes into consideration geographic space and geoinformation technology (Čeh 2003).

In our proposal geoinformation technology is presented with Ontology of GIT and definitions of concepts are adopted from the groups of standards ISO reference model (19101), images and grid data (19121), feature catalogue (19110), coordinate reference system (19111) and spatial schema (19107) and group of CEN standard 12160:1997 (spatial schema, geometric and structural elements – topology, location, geographic identifications – no coordinate links).

Linkage of typological spatial concepts of two semantic models (UOGS and OGIT) is provided through two concepts of UOGS defined as:

- Activity of naming geographic features with proper names
- Activity of establishing geometric schemas in geographic space.

The group of subordinated concepts point, line, area, cell of mosaic, represents the link of basic concepts of geographic space to technological concepts of geographic information systems. Semantic integration of descriptive definitions of spatial locations in no coordinate spatial reference systems is based on the concept of UOGS named "Activity of naming geographic features with proper names".

Semantic GIS and Universal Ontology of Geographic Space

Data collection remains the most expensive and time-consuming part of most GIS projects. It can require up to 60% or sometimes even up to 80% of available time and funds. The discrepancy between high costs of data collection on one hand and low rate of data usage on the other hand poses a need for an approach that will change this situation. GIS should form a support to the activities the humans are participating in. As it is now, GISs are being set up with only a little regard to the specifics of the environment in which they are used (too often a GIS is only a passive model abstracting real world's phenomenon). In order to make geographic information more easily accessible (in an understandable way to everyone – in the terms the non experts are familiar with) a methodology for creation of an ontology focused on human activities was proposed (Čeh, 2003). Universal ontology of geographic space (UOGS) is a semantic reference system that consists of 588 concepts logically organized into 5 strict hierarchic levels. Set of concepts on the lower levels is not exhaustible but upgradeable. Combining the simple ones supports the complex concepts. Some examples of the third level classes of activities are:

- activity of residence,
- activity of acquiring natural resources, energy,
- activity of products manufacturing,
- leisure activities...

The next few lines outline an idea of accessing spatial data with the help of designed "Universal ontology of geographic space" and supportive mechanism of semantic comparisons of available and requested data. First, a user selects a concept (a theme) within the displayed semantic reference system UOGS that she/he is interested in. This can be, for example, a concept of "school". For the next step, user constrains the search to the geographically defined administrative region (state, region, and city). As a result, the semantic part of GIS engine processes the request and proposes a list of all available (distributed) geodata, semantically and geographically matching the concept "school". Assuming that distributed geodatasets offer semantic information which can be mapped by semantic GIS engine, and the fact, that used GIS – semantic GIS engine – uses "Universal ontology of geographic space", the proposed list consists of semantically referenced geodata sets, describing phenomenon/objects that are the cause/consequence of the specified concept. Next, the user can read metadata about the quality of each proposed geodata set; compares the data quality estimates and picks up (buys) the one that best fits her/his needs.

Hierarchies that tend to expand exponentially with depth can be laid out uniformly in hyperbolic space, so the distance between parents, children, and siblings is roughly the same across the hierarchy (Lamping, Rao, 1996). Strict hierarchic structure of UOGS is necessary to establish the semantic measures. Therefore we propose the use of hyperbolic browser (fisheye) scheme to support visualizing and manipulating large hierarchies of UOGS to the non-GIS experts (Figure 2).

GEODATA USABILITY

Geodata presents an integral part of each GIS. In cases when geodata are of low quality, the information generated by the GIS is of poor quality as well. The GIS in that way produces and fosters the transmission of the information of questionable quality. Additional costs and even damage may be a result when important environmental or any other decisions are based on such kind of information. In this respect, it is important to utilize elaborated but economic data acquiring technologies which are able to provide quality up-to-date geodata.

Remote sensing, aerial photogrammetry and GPS technology have become traditional spatial data acquisition methods by now. Probably every GIS product handles vector data and supports manipulation of georeferenced raster maps (photographs). However, the development of new modern but mature spatial data acquiring techniques has raised attention of the GIS vendors in recent years. Still, these techniques are characterised as unconventional.

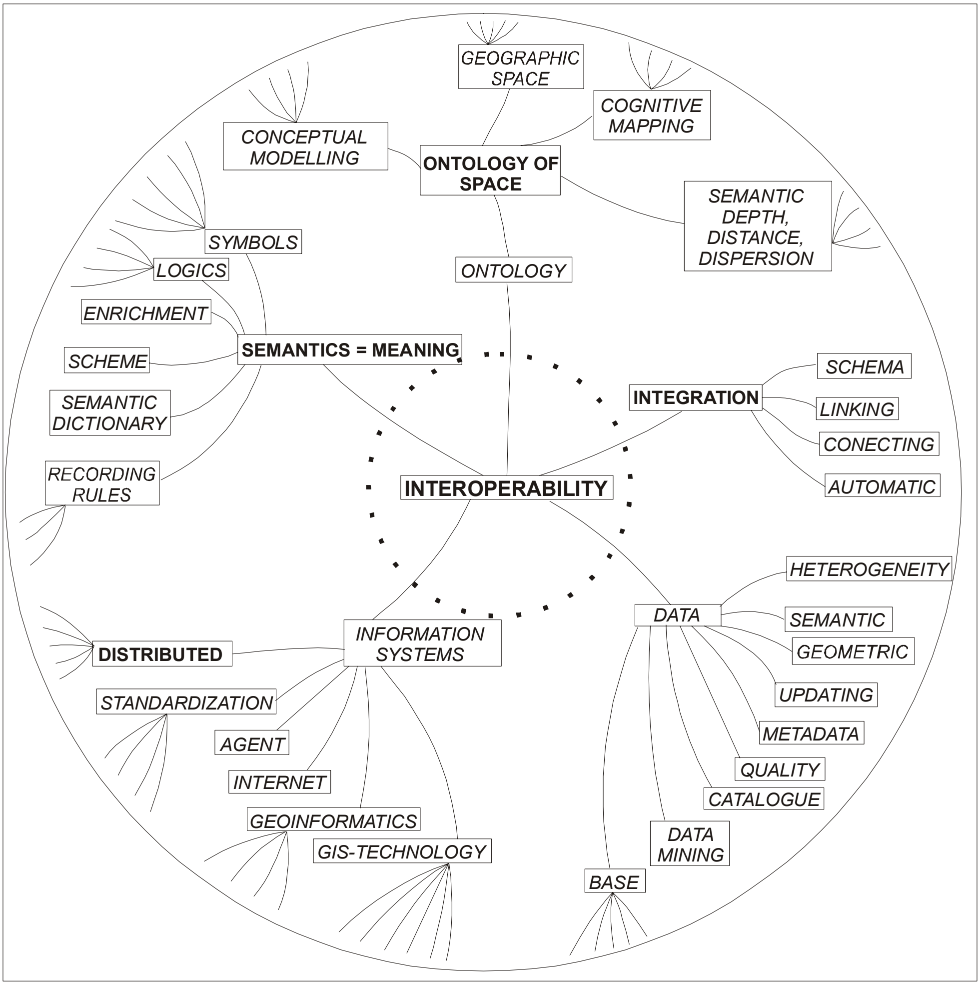

Figure 2: Display of basic terms of interoperability with hyperbolic browser which supports dynamic display of hierarchic models on hyperbolic plane.

Technology consisting of remotely-controlled helicopter/balloon equipped with (photogrammetric) digital camera being able to acquire data (photos) in short time at reasonably low costs certainly promises a lot. The same goes for mobile mapping systems. All kinds of data on road equipment have never been easier to acquire with such accuracy before. Aerial and terrestrial laser scanning are emerging and in the combination with photogrammetry very promising technology too. Last, but not least, underwater photogrammetry is another not yet widely utilized but certainly very interesting data acquiring approach.

Since GIS tools depend on the type of geodata that is to be manipulated, it is obvious that all characteristics of the geodata have to be taken under consideration when developing GIS software modules. Today, to be competitive, it seems that making sound partnerships between GIS vendors and companies that master versatile modern spatial data acquiring methods is a must. However, acquiring geodata that is stored, managed and visualised in three dimensions (3D) seems that needs to be encouraged in the future. One of the main reasons is that we live in a 4D world and GIS is expected to provide at least 3 of them. Beside that, humans are more accustomed to percept things from the “ground perspective” than from the “aerial perspective”. In this respect, navigation through 3D presentation of a certain geographic area seems much more natural than navigating the situation in 2D only. However, 2D presentation is better when trying to get an overall view of (extended) geographic areas.

CONCLUSIONS

Time and location are becoming less important in the sense of accessing the information. There is a high possibility, that anyone having at least some experience in using a Web browser and navigating the Web comes upon some sort of geographic data and tools that enable performing GIS operations such as locating, navigating, searching a particular or finding the nearest facility or event. Internet has contributed the most to spatial data and GIS popularisation within distributed computing environments and mobile computing domain. However, that does not mean, that desktop GIS applications will disappear in the future. Sound desktop GIS solutions will be needed for performing complex spatial analyses involving large amount of raster, vector and semantic geodata.

To ensure the global use of the GIS technology, certain things will have to be improved. The possibility of abusing sensitive and confidential data should be minimized to the greatest possible extend. Next, providing consistent and secure GIS services payment transactions is necessary to generate revenue to vendors and trust in users. And the most important issue, offering user-friendly and easy-to-use GIS services supported with rich spatially and semantically enabled content will be of vital importance. Considering this, enhanced activities addressing current semantic non-interoperability will play an important role in the future.

The aim of our research group is to contribute to the opened issues of semantic interoperability and data integration with a unique approach to the conceptual modelling of geographic space through Universal ontology of geographic space, defining this way a semantic reference system and designing a process of semantic-geographic search within distributed geodata environment. One of the topics of our further research is the implementation, deployment and testing of the described geodata search approaches in the distributed environment, namely the Internet, which may according to the expectations, simplify searching and accessing the geodata to a great extent.

BIBLIOGRAPHY

Burrough, P. A., McDonnell, R. A., Principles of geographical information systems. Oxford University Press, Oxford, 1998.

Čeh, M., Semantic integration of spatial databases/Semantična integracija zbirk prostorskih podatkov, ZRC Publishing / Založba ZRC, Scientific Research Centre of the Slovenian Academy of Sciences and Arts, Ljubljana, Slovenia, 2003.

Kuhn, W., Raubal, M., Implementing semantic reference systems. 6thAGILE, Lyon, France, 2003.

Limp, F., User needs drive Web mapping product selection. GEOEurope, (3), pp. 40-46, 2001.

Lamping, J., Rao, R., Visualizing Large Trees Using the Hyperbolic Browser. Annual conference on Human-Computer Interaction, CHI 96, Association for Computing Machinery, 1996.

http://siglink.acm.org/sigchi/chi96/proceedings/videos/Lamping/hb-video.html (accessed 10. Feb. 2004),

Plewe, B., A. simple Web mapping solution for complex spatial databases. http://www.svgopen.org/abstracts/plewe__webmapping_with_spatial_databases.html (accessed 18. Oct. 2002), 2002.

Quartararo, A. J., Spatial ASP-Internet GIS is Here -Part 2, 2002. http://www.directionsmag.com/article.php?article_id=147 (accessed 15. Oct. 2002)