United Nations

E/CONF.93/L.12

Economic and Social Council

Distr.: Limited
30 November 2000

English only

Seventh United Nations Regional Cartographic Conference for the Americas
New York, 22-26 January 2001
Item 7 (a) of the provisional agenda*
Reports on achievements in surveying, mapping and charting in addressing national, subregional, regional and global issues, including policy and institutional issues

Education and the geospatial data infrastructue A project based approach

Paper submitted by the Netherlands**

* E/CONF.93/1.

** Prepared by Professor Menno-Jan Kraak, International Institute for Aerospace Survey and Earth Science, Division of Geoinformatics, Cartography and Visualization.

00-77873 (E) 081200

1. Setting

Who needs to know about a geospatial data infrastructure (GDI)? It will be obvious that any educational program that deals with geospatial data has to cover the topic of GDI. This is irrespective of the emphasis of such a program. Courses on mapping, on surveying, on remote sensing, or those that claim to deal with geoinformatics, and even those dealing with applications, none of them can ignore GDI.

This paper describes a project-based approach to the GDI. The project's objective is to create a realistic working environment for the geospatial data handling processes to be used in education, research and consulting. The project is an initiative of ITC's geoinformatics cluster, which will provide sample material for demonstrating a GDI that can (simulate) real world problems. The project focuses on integration and processing of geospatial data from various sources for the municipality of Enschede, ITC's home base. The practical approach will guarantee a direct link between academic theory and the practicalities of all day life. It will strengthen ITC's contacts both the policy and practical levels with those organisations dealing the real world geoinformatics problems, and while executing the project both staff and students will learn.
But what is a GDI? According to Groot & McLaughlin (Groot and McLaughlin, 2000) "it encompasses the networked geospatial databases and handling facilities, the complex of institutional, organisational, technological, human, and economic resources which interact with one another and underpin the design, implementation, and maintenance of mechanisms facilitating the sharing, access to, and responsible use of geospatial data at an affordable cost for a specific application domain or enterprise". How to teach this complex reality?

The GDI is treated in the framework of the geoinformatics and geoinformatic management programs. At ITC Geoinformatics is defined as: "... is the combination of several methods and techniques for the acquisition, storage and processing of geospatial data, as well as the dissemination of these data and (geo)information services". This implies: the development and application of concepts for geoinformation modelling, for information extraction from data, as well as information sharing and presentation, the development and implementation of concepts for the institutional setting, organisation, structuring and management of above processes.

Reading the above definition allows different views on geoinformatics, and within each view one can emphasize different perspectives. Here we consider two of those perspectives: the first concentrates on the structure of the geospatial data handling processes and the second on the context from which one looks at the data handling. Both are combined in figure 1.

The horizontal axis in the diagram represents the first perspective and distinguishes between different phases in the process: data acquisition, storage and query, processing and presentation, and dissemination and use. The diagrams vertical axis distinguishes the different perspectives on the data handling process: application, technology, information management, and institutional setting and policy. Examples of applications could be cadastre and land administration, water management, natural resources management, sustainable land use etc.

Proces / *Context*	**Data acquisition**	**Storage & query**	**processing & resentation**	**Dissemination & use**
Application				
Technology				
Information management				
Institutionel setting & policy				

Figuur 1. Structure and context of the geospatial data handling process.

The multiple views on geoinformatics put some demands on education. Students should have understanding and insight in geoinformatic technology, but they should also have knowledge of the application fields where the geoinformation is used. Especially today the introduction of geoinformation technology has tremendous effect on an organisation's working methods and management. Both internal and external responsibilities regarding information services will change and boundaries between the traditional disciplines involved become vague or disappear. Students will be working in interdisciplinary environments and have to be prepared for this, which puts certain constraints and demands on education.

The GDI-project will be used throughout the educational programs. It will play a major role at the start of the program, illustration the need for a structured approach to geospatial data handling and justifying the program as such. During the remainder of the course it will be used as example and illustration in many practical exercises while individual assignments and MSc-research topics will be executed in the framework of the project.

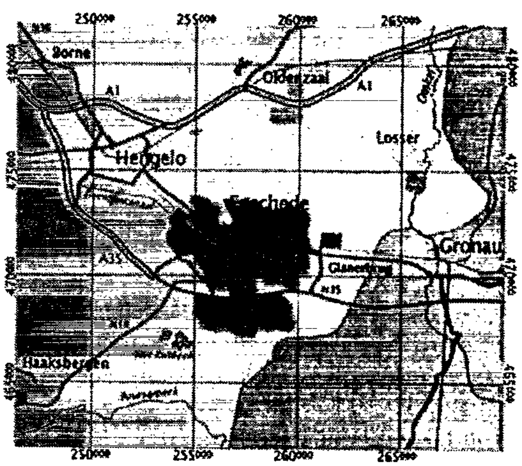
2. The background of the project-based approach

In the past ITC's geoinformatics programs were closely linked with the activities of National Mapping Agencies. As recognised, most of these mapping organisations are in the midst of change: a change guided by new technology (e.g. the World Wide Web), by government decisions (cut-backs / privatisation), and competitive market demands (an extension from a purely supply-driven to a supply-&-demand-driven approach). What should educational programs do to be able to react pro-actively to these changes? Thinking of the GDI Infrastructure this brings up the question: "Do the NMA's offer the core data that users of geospatial data need or require?". The question has to be answered since of all data supplied, geospatial core data is the skeleton for the National, if not Global Spatial Data Infrastructure. Traditionally, the National Mapping Agencies were the providers of these data.

"What should be the minimal requirements for the contents, format, structure and accessibility of a topographical (framework) database, which can be applied by a wide range of geospatial data users (e.g. in a GDI-environment)?" The answer to the question of database contents should lead to the distinction of *essential* topographic features (required by any application – the geospatial core data) versus *additional* topographic features (required only for specific applications – the geospatial framework data). This distinction should be supported by a cost-benefit analysis of number of features versus costs related to acquisition and maintenance of those features. Starting point for the determination of contents can be based on already existing standards (as available from FGDC, for example). *Geospatial core data* include geodetic control data (spatial reference system), fundamental topography (used as additional geometric reference represented in the terrain), the digital elevation model, administrative boundaries and postal codes (essential to link socio-economic data to physical data), official geographic names.

The project, in its preliminary phase, concentrates on core- and some framework data (application data) only, since this core data form the base of almost any geospatial application. The development of an efficient GDI also calls for ways of easy data accessibility and data exchange. In this respect, the World Wide Web (WWW) will play a very important role. Communication over the WWW is virtually platform-independent, unrivalled in its capacity to reach many users at minimal costs and easy to update frequently. The WWW allows for a dynamic and interactive dissemination of geospatial data, offering new processing and visualisation techniques, and new use possibilities, especially not seen before with traditional maps and map-based databases. All data sets and externally acquired data sets should be accessible from every workstation – through a suitable infrastructure – for all students and staff and, in addition, need to be described according to standard guidelines in the form of meta-information. Although presented via the WWW not all data can, due to copyrights, be made public.

In a broader perspective the project will define and test the conditions for this kind of information management, by setting up a geo-infrastructure for students and staff. All kinds of case studies, focussed at parts of the production process, will contribute material for education and research. Besides, cooperation with external organisations that produce and use geoinformation (for example, contacts have been established with geospatial data providers), will provide knowledge from practical experience.



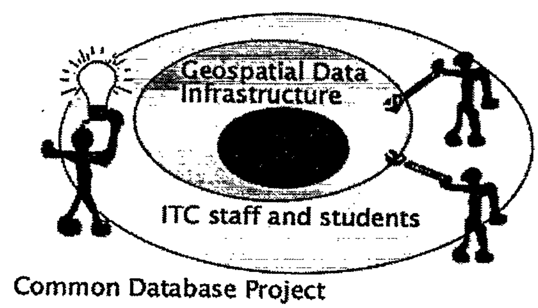


Figure 2. Location of Enschede and setting of the project.

The project will further demonstrate the benefits – in terms of time and finances – of maintaining a common database (and infrastructure) in relation to the conventional, less co-ordinated distribution of geospatial data. The project will focus on a data rich environment – the initial attention is directed towards a compilation of data for the municipality of Enschede and surrounding area. Another advantage of the Enschede environment is that the field is next door, and it contains both large urban and non-urban areas with tension on the border between the two (see figure 2). Furthermore, the choice for Enschede enables intensive cooperation with local authorities, which currently try to tackle similar problem areas in practice. At a later stage, the findings can be matched with data-poor environments in developing countries. Results can be projected on the practical working environment of ITC students, by providing indications for various situations in developing countries. Their employers are often located in less ideal - i.e. data-poor - environments, in which many geospatial data have to be collected and the setting up of a GDI has yet to start. With this experience ITC students will be better prepared to direct the future of their organizations.

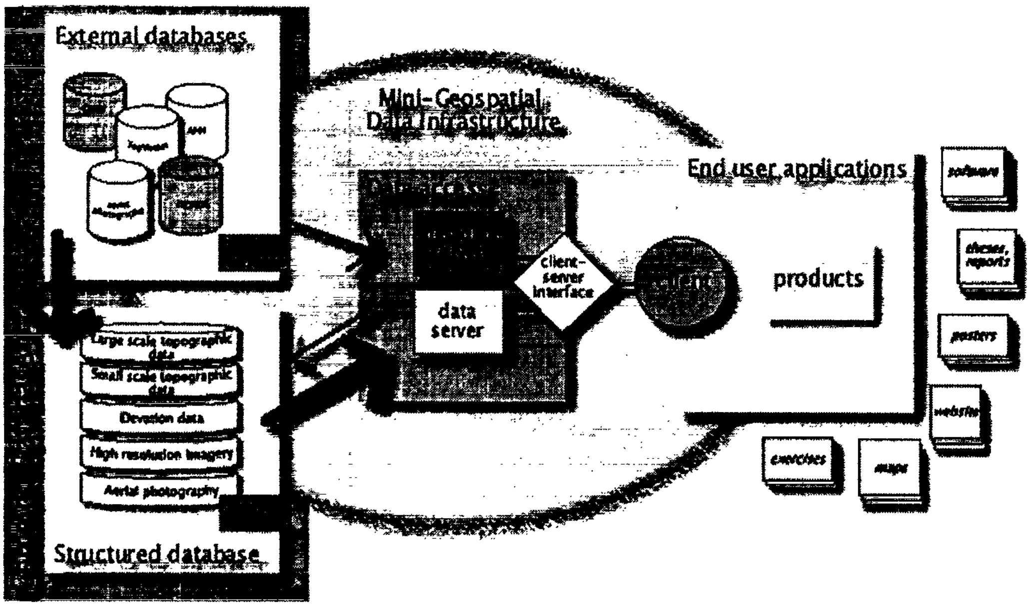


Figure 3. Structure of de mini GDI with a common database, the GDI and end user applications.

3. The project structure

Figure 3's diagram represents the structure of the mini GDI. The concept distinguishes three major components:

- A common database

Which holds two types of data, external data (as provided by data suppliers) and structured data (pre-processed in a way that makes the data accessible for most applications within the ITC working environment). Most often users will make use of the structured data (as the pre-processing stage does not have to be reproduced again), but of course some users may prefer access to the 'original' data for specific applications. Both external and structured data relate to the urban and rural environment of Enschede. Current available data is given in table 1. It is the intention of the project to focus initially on (topographic, cadastral and administrative) geospatial core data; at a later stage, thematic data may be added to the database (to evaluate integration capabilities). Next to collecting data attention is paid to (WWW) interfaces for querying the meta-databases of all available data, as well as for analysing and downloading components of the common database.

- A mini geospatial data infrastructure

The online accessibility of the data should be provided for by the implementation of a mini-geospatial data infrastructure (GDI) that incorporates clearinghouse functionality, options to consult metadata and actual download of data (sub) sets.

- End user applications

Clients can be ITC staff and students who wish to use the data for their expertise-specific applications, leading to all kinds of output (from exercises to reports, from pr-material to journal publications, from maps to websites). Students already have executed several projects, among them Constraint rules for geopatial database updating; A client/server clearinghouse architecture with geospatial data downloading capacity, and GIS application on the Internet for exploring and dynamic visualization of a topographic database. Colleagues from application divisions have used the data to execute a research project for an environmental organisation interested in the effects of some measures taken is a protected nature area in the municipality. The geospatial core data was, together with field data, used to do a geospatial-temporal study.

Table 1. Available data

type of data	provider	year
Large scale topo data 1:1.000	Municipality	1998
Topographic data 1:10.000	Topo Survey	1995
Topographic data 1:50.000	Topo Survey	1995
Aerial Photography 1:10.000	Municipality	1998, 2000
Landsat	NLR	1997, 1999
Ikonos	EON	2000
Height data (Laser-altitude data 1 point every 16 m^2)	Survey Dept	2000
Cadastral boundaries	Cadastre	2000

4. Project use

In the project based approach to the GDI the common database and its infrastructure should serve as an environment to solve general questions that can be formulated with regard to a geospatial data infrastructure in practice. Each of these central questions will address different fields of attention within the process of geoinformation production that we teach to our students and in which we do research. Questions that might arise which each user involved in the project include:

	Issues for CDP
What is in it?	• External and structured data for the municipality of Enschede
How is it managed?	• Management of database structure, updates and legal access restrictions
How do I get it?	• Online access facilities: can be a simple shared directory, to an advanced solution with GIS-functionality on either client or server side.
What can I do with it?	• An unlimited array of applications that demonstrate solutions to problems in geospatial data handling

Answers to these basic questions are given at the accompanying website (www.itc.nl/~cdp). The homepage is also used to inform ITC staff and students, and other interested researchers on the project background, activities and findings. The database(s) will only be made accessible within ITC. Publications will aim at the dissemination of project findings, attracting reactions from a broader audience (users, producers).

The project will develop a theory for the definition of a Common Database (the geospatial core data), in the first place for educational purposes. Such a theory may evolve from a matrix, which explores flexibility requirements by confronting object types with their specifications. These specifications are related to:

- multi-scale characteristics (resolution vs. scale, scale-independency)
- update frequency
- coordinate system (GPS-aspects)
- accuracy (geometry, attribute, topology, etc.)
- 2D, 2.5D, 3D aspects
- time (4D?)
- meta-language

Additionally the view presented in figure 1 will influence this.

The projects to be executed can also be distinguished in applications involving the development of the common database itself ('builder applications'), and applications that need the data provided by the common database ('data end-user applications'). The cover many aspects of the geospatial data handling process and the project blend into the educational program. Actually it becomes the red thread through the program.

Examples of builder applications database update procedures, metadata establishment (+ storage, updating), standards and data quality assessment, optimisation of data storage space and access-time, testing the applicability of standards, distributed database design, client/server architecture between common database and users, design and implementation of interfaces for data providers (as part of update procedures) etc.

Examples of data end-user applications are spatial database management, analytic description of geoinformation production line (from base data to end product, low – high profile solutions), optimisation of geoinformation production lines (processes and workflows) in terms of testing performances of alternatives (flexibility/product diversity, quality, time and costs), cartographic tools for automated generation of graphic representation model from the digital landscape model for different presentation media (from hard copy, electronic to WWW) and forms (static vs dynamic presentation have different requirements, integrated mapping projects, etc.

5. Conclusions

A project-based approach to the Geospatial data infrastructure will be and interesting and successful formula in geoinformatics education to teach and learn. It offers a framework for both pragmatic as well as theoretical perspective. It confronts both staff and students with the day-to-day problems. It will stimulate the relations between data providers, educational institutions and the user community of geospatial data. The project can potentially be the red thread through the program linking the many seemingly 'isolated' items together.

References

Groot, D. and McLaughlin, J., (Eds) (2000). *Geospatial data infrastructure - concepts, cases, and good practice.* Oxford: Oxford University Press, pp: 286; ISBN 0 19 823381 7.