ITC

Provisional Plan for Common Database Project (CDP)

Contents:

1. Aim of Common Database Project

Topographic base data form the framework of almost any spatial application. Currently available topographic databases often contain an insufficient range of fundamental features, whereas other additional features may be regarded as less essential or even superfluous. The offered data structure does not always serve direct application in a GIS, since this is mostly determined by the main interests of the data producing organisation (e.g. data sets structured for cartographic purposes in the case of the National Mapping Organisations).

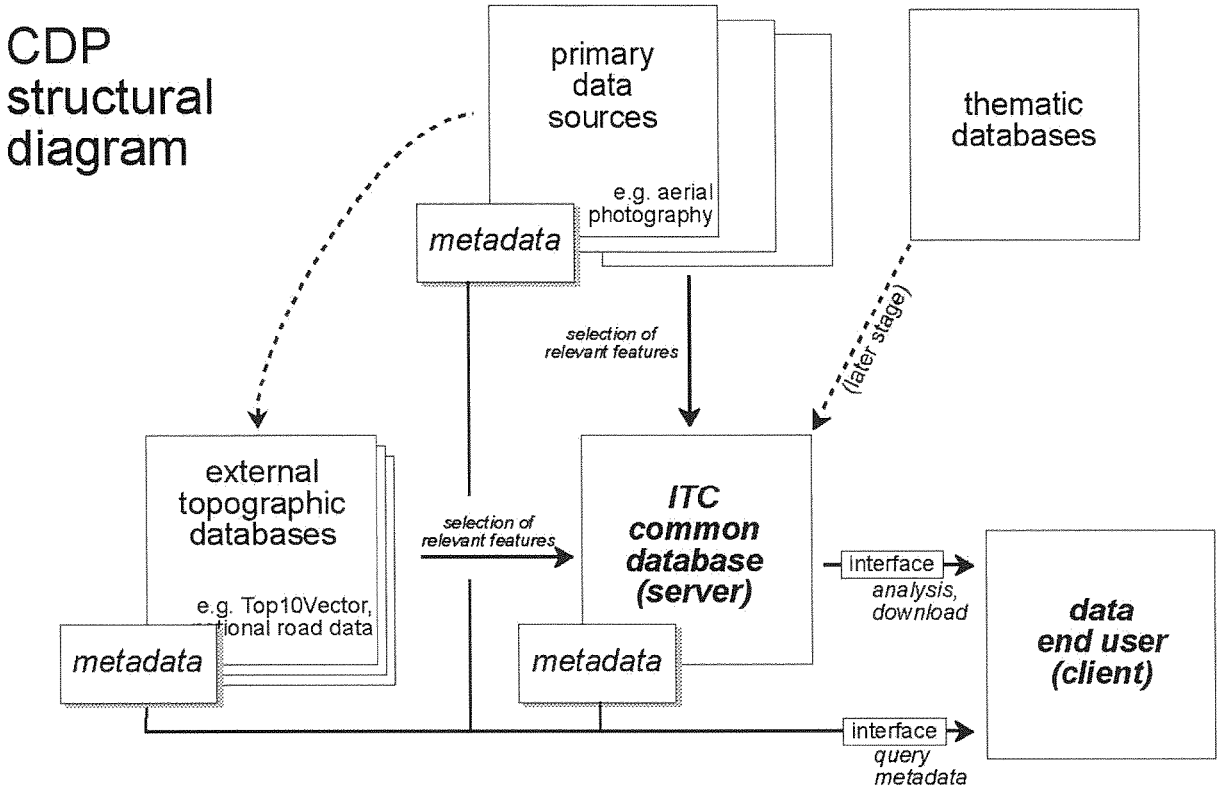
This project will define and compile the most suitable set of base data, in terms of data features, data format, data structure and data accessibility, usable for a wide range of spatial applications. This set of base data will be referred to as the Common Database.

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. At a later stage, the findings can be matched with data-poor environments in developing countries.

On the whole, the project aims to create a data set to test and demonstrate the integrated approach of the Geoinformatics Divisions to the spatial data handling process. The data set will first serve as material to organise final projects and MSc-studies, but later also for support in research and consulting applications.

The project is a joint initiative of the Geoinformatics Divisions at ITC:

1. Division Geoinformatics and Spatial Data Acquisition (contact: G.C. Huurneman)
2. Division Geoinformatics, Cartography and Visualisation (contact: R.M. Hootsmans, chair)
3. Division Geoinformatics, Spatial Information Theory and Applied Computer Science (contact: R. Lemmens)
4. Division Geoinformatics Management and Infrastructure (contact: C.M.J. Paresi)



2. Central question

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 spatial data users?

The answer to the question of database contents should lead to the distinction of *essential* topographic features (required by any application) versus *additional* topographic features (required only for specific applications). 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).

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 spatial data.

3. Some background

(cf. Kraak, 1998. National Mapping Organisations and the World Wide Web. *Geoinformatics*, vol. 1, no. 7, pp. 6-7)

Of all data supplied, spatial base data is the framework for the National, if not Global Spatial Data Infrastructure. Traditionally, the National Mapping Organisations were the providers of these framework data. Most 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!),

and competitive market demands (an extension from a purely supply-driven to a supply-&-demand-driven approach).

Thinking of the Spatial Data Infrastructure brings up the question: "Do the NMO's offer the framework data that users of spatial data need or require?"

For instance, one can look at a product on offer with most European NMO's, the 1:50,000 database. It should be realised that most often the contents of this database has been designed for paper map series in the days of Napoleon. Much of this content is still very valid: data on the infrastructure and land use are examples. However, also many Napoleontic artefacts, for instance, registration of objects that supposed to be barriers for the foot soldier can still be found. On Dutch topographic maps some classifications used still hold samples, such as canals wider than 3 m, fences higher than 2 m, and escarpments higher then 1.5 m. Many objects are recorded in a way that supports graphic representation purposes, but can seriously inhibit GIS-analysis and -processing. This may not exactly be what we need with respect to the contents of a modern framework database within a National Spatial Data Infrastructure.

The development of an efficient Spatial Data Infrastructure also calls for ways of easy data accessibility and data exchange. In this respect, the World Wide Web (WWW) is likely to 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 spatial data, offering new processing and visualisation techniques, and new use possibilities, especially not seen before with traditional maps and map-based databases.

4. Initial project activities

This section briefly summarises the initial activities only. The project will likely evolve into multiple (but related) directions. New activities can be defined at a later stage in the project. The activities are listed in an arbitrary order: they can be deployed simultaneously.

4.1. Inventory of user and producer requirements

The current 'state of the art' within the working environment of both users and producers of topographical framework data will be evaluated, in order to obtain a starting-point for the definition of the common database. Topics for such an inventory are:

- existing framework data standards (FGDC – [www.fgdc.gov/framework/frameworkintroguide/], CEN TC287 [website Eurogi], etc.)
- Municipality of Enschede (via Chris Paresi/Ben Draaijers)
- Province of Overijssel (Provinciale Staten, Zwolle/ via PhD-research Corné van Elzakker)
- Water Board "Regge en Dinkel" (via Dick Groot/Bert Raidt)
- Meetkundige Dienst, Rijkwaterstaat (via Lucas Janssen)
- Rijkswaterstaat/ANWB: Nationaal Wegenbestand (via Lex Polderman)
- Dutch Topographical Service (via Paul van Asperen)
- Ordnance Survey (specifications listed on the Internet)
- Swiss Topographical Survey (info via Willy Kock)

- IGN France, Belgium
- *etc.*

4.2. Acquisition of primary data sources

It is the intention of the project to focus initially on (topographic, cadastral and administrative) framework data; at a later stage, thematic data may be added to the database (to evaluate integration capabilities).

This stage will concern setting up of a software/application independent database of Enschede, which will then be used for the compilation of the common database. This includes (WWW-)interfaces for querying the meta-databases of all available data, as well as for analysing and downloading components of the common database.

The selection of primary data will mainly be demand driven – to be determined by the objectives of the common database. However, the collection of data which can be regarded relevant has already started. Possible data to be used as primary data source for the common database:

- Top10Vector and Top50Vector databases from the Dutch Topographical Service (on condition that this database will not be transferred to third parties), *available.*
- Large scale base map 1:1,000 (GBKN), *not yet available.*
- Laser-altitude data from the Meetkundige Dienst (Ministry of Transport and Water Management), *not yet available.*
- Aerial (ortho-)photography, remote sensing images: an inventory of images present at ITC is necessary (date of data capturing!), *not yet or partly available.*
- Cadastral boundaries: parcel outlines (Kadaster, Dutch Cadastral Service), *not yet available.*
- Administrative boundaries: neighbourhoods, districts (Municipality of Enschede), *not yet available.*
- Hydrology: waterways (Waterschap Regge en Dinkel, Water Board), *not yet available.*
- *etc.*

4.3. Theoretical framework

The project will develop a theory for the definition of a Common Database, in 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 should include, for example:

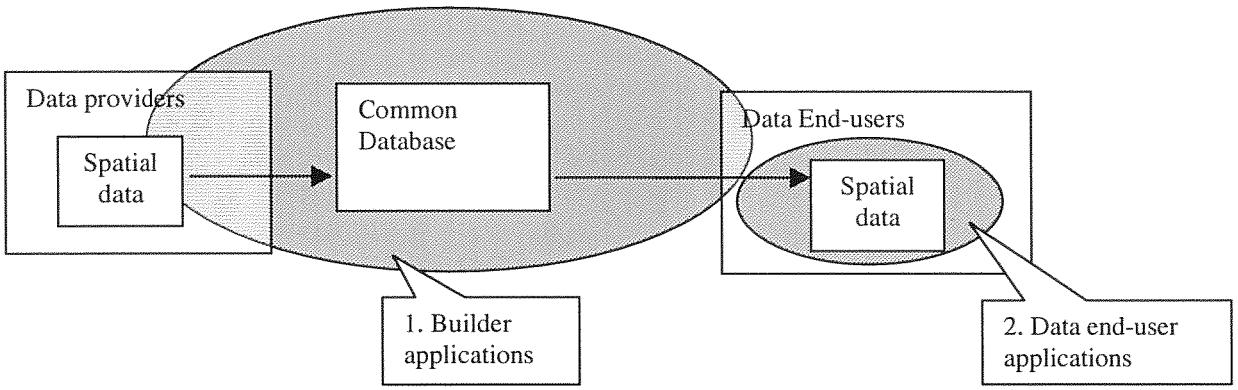
- multi-scale characteristics (resolution vs. scale, scale-independency)
- update frequency
- coordinate system (GPS-aspects)
- accuracy (geometry, attribute, topology, etc.)
- 2D, 2.5D, 3D aspects
- time (4D?)
- meta-language

A third dimension within this matrix, besides object types and specifications, can consider a range of applications for educational (or research, consulting) purposes.

4.4. Sample applications

Each division within Geoinformatics can put emphasis on its specialisation, for example GIS-analysis (utility management, civil engineering) and data visualisation (cartography), by means of project teams (which can also act 'inter-divisional'). The lists of samples mentioned in this section is still subject to changes.

A distinction is made (see figure) between 1. applications involving the development of the common database itself ('builder applications'), and 2. applications that need the data provided by the common database ('data end-user applications').



4.4.1. 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 (CEN-TC 287, ISO-TC-211)
- distributed database design
- client/server architecture between common database and users:
  - thin client / heavy server vs. thick client / light server implementation issues
  - design and implementation of intranet user-interfaces for data end-users
  - design and implementation of interfaces for data providers (as part of update procedures)
  - implementation of data format conversion procedures
- *etc.*

4.4.2. Data end-user applications

- spatial database management
- GIS network analysis
- 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)
- metadata communication: consulting and unambiguous representation (e.g. does visualisation of spatial metadata require other rules than regular data visualisation?)

- multiscale and multitemporal applications, scale independent applications: implications for database design and management (also in relation with representation and visualisation requirements)
- 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)
- large scale network database (e.g. sewage system) enrichment with laser height measurements
- improvement of classification of raster image data
- integrated mapping projects
- parcel based mapping and analysis for cadastral procedures
- analysis, design and implementation of cadastral and municipal information systems
- use of land survey procedures (conventional/GPS) for database updates
- research on spatial information theory (e.g. topological issues, spatial data structures)
- *etc.*

4.5. Data accessibility and publications

The project will maintain a homepage at the ITC-website (http://www.itc.nl/~cdp) 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, for known ITC-staff via the Intranet. Publications will aim at the dissemination of project findings, attracting reactions from a broader audience (users, producers).

5. Initial time and budget planning

On short term, a staff and budget planning has to be made to embed the project parallel to or even within on-going activities of the Geoinformatics Divisions, and to formally enable the project's foundation.

In the first year, the building of the Common Database will be executed in the framework of preparing education material for the Geoinformatics courses. A great part of the activities will also be dedicated to setting up proper client/server interfaces for the consulting of metadata and the analysis/downloading of components from the common database.

After this period, the database will be expanded to other disciplines. External funding will be needed to finance maintenance and education.