Harmonization of geo-information related to the lifecycle of civil engineering objects - with focus on uncertainty and quality of surveyed data and derived real world representations

PhD Research Proposal

Wiebke Tegtmeier, MSc

GISt Report No. 45

March 2007

GIS TECHNOLOGY

TU Delft

OTB Research Institute for Housing, Urban and Mobility Studies

Delft University of Technology

Harmonization of geo-information related to the lifecycle of civil engineering objects - with focus on uncertainty and quality of surveyed data and derived real world representations

PhD Research Proposal

Wiebke Tegtmeier, MSc

GISt Report No. 45 March 2007

Summary

In infrastructural projects, communication between involved parties is difficult. This is, among other things, caused by a lack of quality and uncertainty information concerning collected data and derived real world representations. Particularly in subsurface geotechnical representations uncertainties are high, since only sparse information is available for the interpretation. This leads to the introduction of "interpretational uncertainties" into the representation; that are, uncertainties introduced by the expert using own knowledge and experience for the data interpretation. That is what, in addition to the variety of data and information types, makes a harmonization of geo-information extremely difficult. This report summarizes available methods and software packages as used by different professionals in infrastructural development for the representation of real world and design objects as well as for the management of geo-information. Furthermore, it emphasizes existing problems and gaps towards the harmonized handling of geo-information including quality and uncertainty estimations for the different real world representations.

ISBN: 978-90-77029-20-6
ISSN: 1569-0245

Section GIS technology
OTB Research Institute for Housing, Urban and Mobility Studies
TU Delft
Jaffalaan 9, 2628 BX Delft, the Netherlands
Tel.: +31 (0)15 278 4548; Fax +31 (0)15-278 2745
Websites: http://www.otb.tudelft.nl
http://www.gdmc.nl

E-mail: w.tegtmeier@tudelft.nl

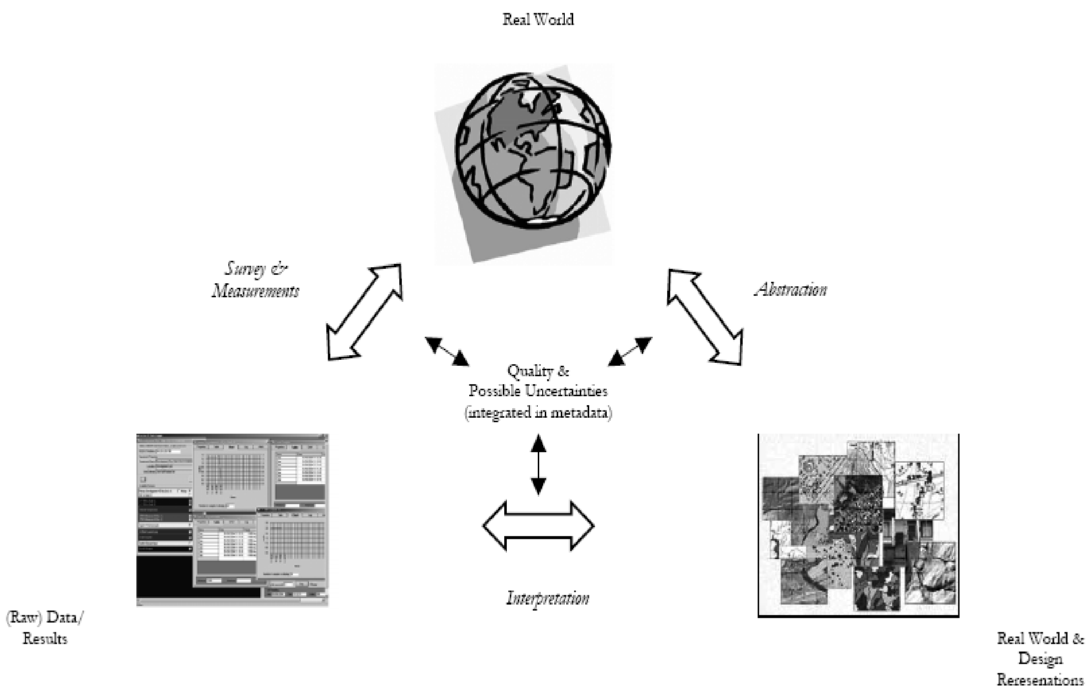
Extended Summary

Around the world people are busy with the realization of infrastructural projects. Different tasks must be accomplished – infrastructural projects are planned, designs changed, existing structures maintained or abandoned, etc. – all asking the skills of a variety of professionals. In order to fulfill these tasks, a number of different representations are developed by the different experts to achieve a picture of the real world situation at the construction site. In general terms, it can thereby be differentiated between three basic types of representations that are namely:

1. Subsurface real world representations
2. Surface (and above) existing real world representations
3. Design real world representations

For each of these representations, different methods are applied for the collection of data and the development of the representations. In addition to the data and information used to create the real world representations, some of the data sets and representations of the data contain information about their quality and possible uncertainties; however, this is not always the case, although it would be most desirable to include estimations of quality and possible uncertainties in the metadata ('data about data') of the various real world representation.

In order to get an understanding of the problem that we are facing and trying to solve with this research, the relationships between the real world, its representations, the collected data and quality and uncertainty information must be clear.



The main problem in large civil engineering projects is, thus, that for the creation of each of the three types of representations (see above) large quantities of geo-information are generated and different types of geotechnical and geological data (e.g. GIS-, CAD-, and other data sets) collected as well as (re-) used, however, usually without any (or with insufficient) indication regarding the uncertainty of the information.

This lack of quality/ uncertainty information, and also the use of different types of data structures and geo-information management systems are the main obstacles when trying to achieve data harmonization in large infrastructural projects. This, however, would significantly facilitate the (re-) use of geo-information as well as the co-operation and communication between companies involved in the same project. The question is thus: How can geo-information be harmonized and equipped with quality/ uncertainty estimations? Still, a proper solution for this question is to be seeked for and this research will approach a solution to this question.
What would already improve the communication between the parties involved in infrastructural development is the use of common routines and schemes. Regarding the collection of the data, the development of real world representations as well as the exchange of data and representations a number of standards (e.g. NEN, ISO, OGC) are available that are defining, for example, the set-up and content of various types of metadata or the spatial and temporal schema of geographic information. By using these standards, a common way of working and easier communication can be achieved. Frequently used standards are, for example:

- ISO 19107:2003 Geographic Information – Spatial schema
- ISO 19108:2003 Geographic Information – Temporal schema
- ISO 19113:2002 Geographic Information – Quality Principles
- ISO 19114:2003 Geographic Information – Quality Evaluation Procedures
- ISO 19115:2003 Geographic Information – Metadata
- ISO 19123:2005 Geographic Information – Schema for coverage geometry and functions
- NEN 3610 Basic scheme for geo-information – Terms, definitions, relations and general rules for the interchange of spatial information of spatial objects related to the earth's surface
- OGC 05-087r4 Observations and Measurements - Framework for measurements and encoding

Regarding the scope of this work, this research, which is part of the Dutch Bsik RGI project GIMCIW ('Geo-Information Management for Civil-Engineering Infrastructure'), can be considered to be in one line with the goal of the section GIS technology at the research institute OTB of the TU Delft. This overall goal of research is 'to provide/develop the technology, including the knowledge behind it, to stimulate the realization of the geo-information infrastructure (GII)'. With it, multiple partners in infrastructural development use and share information provided in a common network. The engine(s) of the GII nodes throughout the world will be the Geo- Data Base Management Systems (DBMSs) filled with geographic data as well as various web services, allowing the data exchange via the internet and different portals. This research is thus not isolated, but fits within the overall geo-information science research agenda, which has been developed in the Netherlands in form of the Bsik knowledge project proposal 'Space for geo-information' (in Dutch: 'Ruimte voor Geo-informatie (RGI)'). Furthermore, this work is closely related to the work of several (international) initiatives dealing with the standardization and harmonization of geo-information, such as INSPIRE or HUMBOLDT. These initiatives are, however, rather focused on domain related semantics standards dealing with different specializations at a time (e.g. geology, topography, cadastre, etc.).

Contents

1 Introduction

Around the world people are busy with the realization of infrastructural projects. Different tasks must be accomplished – infrastructural projects planned, designs changed, existing structures maintained or abandoned, etc. – all asking the skills of a number of professionals. With it, various problems need to be tackled, often requiring the combination of a variety of data and knowledge as collected by the different professionals (e.g. civil engineers, engineering geologists, GIS technologists, etc.) involved in the project.

The main problem in these large infrastructural projects, as known today, is the difficulty regarding data harmonization. This is, among other things, often caused by uncertainties concerning data as well as real world representations.

Generally, projects for the development of large civil-engineering constructions are characterized by a long lifetime (i.e. tens of years). This lifetime can be subdivided into six main stages (see also Figure 1.1) that are namely:

4. Exploration
5. Planning
6. Design
7. Realization
8. Maintenance
9. Abandon

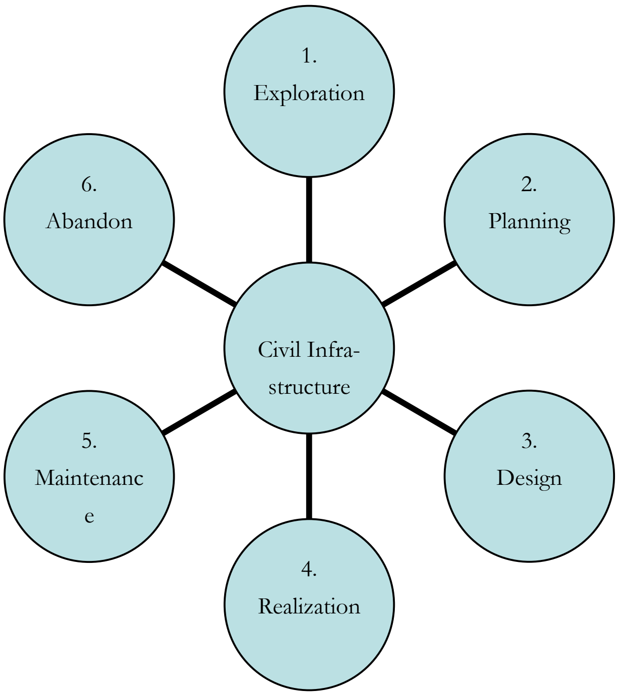


Figure 1.1 Lifetime of civil infrastructures.

Thereby, large quantities of geo-information are generated and different types of geotechnical and geological data (e.g. GIS-, CAD-, and other data sets) are collected as well as (re-) used, usually without any (or with insufficient) indication regarding the uncertainty of the information. This, and also the use of different types of data structures and geo-information management systems are the main obstacles when trying to achieve data harmonization in large infrastructural projects. The question is: How can geo-information be harmonized and equipped with uncertainty estimations? Still, a proper solution for this question is to be sought for. This research will approach a solution to this question.

Regarding the scope of this research project, it can be considered to be in one line with the goal of the section GIS technology at the research institute OTB of the TU Delft. This overall goal of research is 'to provide/develop the technology, including the knowledge behind it, to stimulate the realization of the geo-information

infrastructure (GII)[1]' (Oosterom, 2003). As described by Oosterom (2003) in his research agenda for the TU Delft in 2003, the engine(s) of the GII nodes throughout the world will be the Geo- Data Base Management Systems (DBMSs) filled with geographic data. This TU Delft/GISt research agenda is thus not isolated, but fits within the overall geo-information science research agenda, which has been developed in the Netherlands in form of the Bsik[2] knowledge project proposal 'Space for geo-information' (in Dutch: 'Ruimte voor Geo-informatie (RGI)') as well as across Europe in the INSPIRE[3] initiative.
The research objective of the section GIS technology is thus to support the realization of the Geo-Information Infrastructure (GII) in a technological sense. Part of this will be realized throughout this research, where a harmonization of geo-information (and the implementation of uncertainty indications) is aimed at. The target is to manage the step from implicit semantics to explicit knowledge, which, in the past, was not necessary within a certain domain as the limited number of users had (more or less) the same interpretation of the basic geo-information within their domain. Nowadays, the GII will, according to Oosterom (2003), expand the user community of basic geo-data sets to persons not familiar with the meaning of this information. Proper metadata is therefore necessary to ensure that these users know where they can find what kind of geo-information. As a result, harmonization of semantics will be approached in this research. Furthermore, additional information needs to be attached to any real world representation used throughout an infrastructural project; preferably equipped with uncertainty information for each of the representations. This should significantly facilitate the understanding of the real world representations and finally also the communication between the various parties involved in such projects. By attaching explaining information as well as uncertainty indications in the metadata, it is also tried to motivate the different parties to (re-)use already existing geo-information and real world representations in future projects.

Overview

Following the introduction in chapter 1, chapter 2 of this proposal summarizes the essential background for this research project. First, the main sources of uncertainty will be given and the type of uncertainty, which is of main interest for this research, the so-called "interpretation uncertainty", will be introduced. Also, a variety of software packages frequently used by different professionals in infrastructural projects are introduced.
Chapter 3 is then used to define the main research question as well as related sub-questions and to determine possible hypotheses regarding research and also sub-questions as proposed for this project.
In chapter 4, the research methodology, which is assumed to be followed throughout this research, is further defined. This determines the sub-division of the research project and the different steps that are expected to lead to a solution of the problem.
Chapter 5 is then used to indicate the proposed output as well as various conferences, which might be attended in the course of this research.
In addition, work and budget plans are shown in chapter 6 and 7 as well as a plan of the organization of the project in chapter 8. Finally, the references used for the set-up of this proposal are indicated and a glossary added, including definitions of the

[1] Internationally the GII is sometimes also called the SDI (Spatial Data Infrastructure).
[2] Bsik is an acronym of the Dutch 'Besluit subsidies investeringen kennisinfrastructuur'; in English 'Decision State aided investments knowledge infrastructure'.
[3] INSPIRE – Infrastructure for Spatial Information in Europe

key terms used in this research proposal in order to prevent confusion about the meaning of the various terms.

2 Background

Throughout the various stages of infrastructural projects (see Figure 1) different requirements have to be met by the geo-information or the systems maintaining geo-information. Various parties that are involved in these infrastructural works are all using their own methodology for data collection and management. Unfortunately, availability as well as (re-) usability of different types of geo-information (e.g. GIS, CAD, geotechnical and geological multi-dimensional representations, etc.) is still rather difficult. The variety of representations, data structures and data types make the access and harmonization of geo- information extremely difficult. However, the main problem regarding harmonization of geo-information is usually caused by uncertainties about the quality of both the data collected and the generated representations, aggravating an overview of the available geo-information.

2.1 The quality aspect as part of the metadata in geo-information

Since many people involved in infrastructural projects make use of available geo-information in order to take decisions, their work strongly relies on the quality of this information. This makes it an important aspect of geo-information and, to allow an effective use of collected data and information, it is, according to Dilo (2006), necessary to know its quality. But, before we can start and determine the quality of various types of geo-information, it is important to understand what the word "quality" actually implies.

Numerous definitions can be found throughout the literature. Harvey and Green (1993), for example, in their pioneering paper explored the nature and usage of quality in relation to higher education, where they conclude that quality is often referred to as a relative concept. In general, there are two senses in which quality is relative. First, quality is relative to the user of the term and the circumstances in which it is invoked. Second, is the 'benchmark' relativism of quality. In some views, quality is seen in terms of absolutes. In other views, quality is judged in terms of absolute thresholds that have to be exceeded to obtain a quality rating. Rather than try to define one notion of quality, Harvey and Green (1993) argued that they could be grouped into five discrete but interrelated ways of thinking about quality. Harvey (1995) provides the following overview of the five categories:

1. The exceptional view of quality sees quality as something special.
2. Quality as perfection sees quality as a consistent or flawless outcome.
3. Quality as fitness for purpose/use sees quality in terms of fulfilling a customer's requirements, needs or desires.
4. Quality as value for money sees quality in terms of return on investment.
5. Quality as transformation is a classic notion of quality that sees it in terms of change from one state to another.

The main definition of quality as used by many engineers and scientists and as defined in the ISO standards (ISO 8402, 1995; ISO 9000:2000, 2001) is a version of quality as fitness for purpose, namely quality as satisfying needs: "Quality: The totality of features and characteristics of a product or service that bear on its ability to satisfy stated or implied needs. Not to be mistaken for 'degree of excellence' or 'fitness for use' that meet only part of the definition." This definition will also be applied throughout this research; however, it should not be forgotten that it only covers part of the definition of quality.

As described by Dilo (2006), several factors may in the end affect data quality and cause imperfections in the data. In general, deficiencies in data quality lead to different kinds of imperfection. According to Smets (1991; 1996), main aspects of imperfect data are *imprecision*, *inconsistency*, and *uncertainty*. Imprecision and inconsistency are properties of the data: either more than one or no world is compatible with the available information, respectively. If, on top of imprecision, we attach weights to the worlds to express our opinion about which might be the actual world, then we are confronted with uncertainty (for more information about 'uncertainty' see also next section). Each of these major categories reveals different nuances of imperfection. In line with Smets, Worboys (1998) distinguishes different kinds of imperfection based on factors causing deficiencies in spatial data quality. Imperfection may for example be *inaccuracy* and *error*: deviation from true values, *incompleteness*: lack of relevant information, or *inconsistency*: conflicts arising from the information (see also Dilo, 2006).

Apparently, it is a complex problem to consider together all factors that can affect the quality of geo-information and real world representations. Different causes influencing the quality of geo-information are treated separately by different theories. In this thesis, however, we particularly deal with uncertainty in geo-information and there, especially in subsurface representations; introduced as a component of the interpretation process.

Uncertainty as part of the quality aspect

Lucieer (2004) pointed out that information on thematic and spatial as well as temporal uncertainty is essential in determining the quality of a classification result. As it can affect further processing steps and even decision making, it is important to understand and quantify uncertainty. In addition to quantification of uncertainty, exploration and communication are essential in uncertainty analysis (Lucieer, 2004). Despite the desire to reduce uncertainty from an end-users and decision-makers perspective, it can never be eliminated (Foody and Atkinson, 2002).

In general terms, uncertainty can be described as a measure of the difference between estimation and reality (e.g. the difference of the thickness of soil layer as determined with CPTs compared to the situation in reality; expressed in percentage). This general description comes close to the definition as used in statistics, where uncertainty is defined as 'the estimated amount or percentage by which an observed or calculated value may differ from the true value'. Similar to the many aspects of imperfect data, different types of uncertainty (e.g. uncertainty to spatial prediction, uncertainty resulting from site investigations/ surveys/ measurements, uncertainty resulting from geological interpretation etc.) can be determined regarding the process of developing real world representation for infrastructural projects. The main problems about uncertainties in the process of (sub-) surface representation are caused by uncertainties in data acquisition, data interpretation as well as the selection of appropriate parameters during the interpretation process.

Nowadays, several methods for the determination of uncertainty associated with spatial prediction, as a component of spatial data quality, are available. In this particular case, the choice of the uncertainty estimation technique is for one depending on the quality and quantity of available data, but also, as described by Zhang and Goodchild (2002), on the type of object, for which the uncertainty needs to be determined. Zhang and Goodchild (2002) distinguish two types of objects, for which uncertainties associated with spatial prediction might occur; namely uncertainties in continuous variables and uncertainties in categorical variables. In order to estimate uncertainties in continuous variables, geostatistical methods, such as kriging, indicator kriging or geostatistical simulations are most frequently used (e.g. Orlic, 1997; Al-Abdalla, 1998). Uncertainties in categorical variables, on the other hand, are mainly estimated using probability-based uncertainty calculations (e.g. Zhang and Goodchild, 2002).

In addition to the uncertainty associated with spatial prediction, there are two other sources of uncertainty that one should constantly be aware of throughout any characterization, although it seems to be difficult to do anything about them. The first concerns the potential for investigation errors (e.g. locational errors or measurement errors caused by wrongly calibrated machines). The second concerns the uncertainty inherent to any interpreted information (i.e. uncertainty introduced by the expert during the interpretation, depending on experience and prior knowledge of interpreter) (Houlding, 1994).

According to Houlding (1994), there is generally little one can do about the potential for investigation error in sample and observation values without comprehensive research into each of the wide variety of investigative techniques in common use. There is a general scarcity of information in this regard: if the research has been done, then the results are by no means freely available. Information on the potential for locational errors is almost as scarce. But, in this case, a considerable amount of information has been published (see e.g. Sides, 1992).

As indicated by Sides (1992) and also Houlding (1994), the potential for error and uncertainty resulting from interpretation of geological features is largely subjective. Up till now, there is no way of incorporating it into a computerized approach unless we are prepared to quantify it ourselves during the interpretation process. Unfortunately, this would essentially require a subjective estimate of the possible variation (in three dimensions) of each geological feature of concern.

Thus, various estimation techniques for the determination of uncertainties (e.g. geostatistical methods) associated with spatial prediction have been studied carefully and are frequently applied in practice. But, in order to solve the problem of uncertainties in (sub-) surface representations completely, more research needs to be undertaken. What is still missing in the study of uncertainty is the influence of the so-called ‘interpretation uncertainties’; that are the uncertainties introduced into the representation by the geology experts themselves. These interpretation uncertainties are, especially in geological (subsurface) representations, a dominant source of uncertainty, which is among other things caused by limitations in data quality and quantity and largely influenced by the knowledge and capability of the interpreter.

With regards to this, the first part of this research will be focussed on the quantification and communication of the so-called interpretation uncertainty; mainly regarding subsurface representations. A first step will hereby be undertaken towards the determination and quantification of uncertainties in subsurface representations. However, the whole problem might not be solved and a complete description of uncertainties concerning the different real world representations not be achieved during this study. But an attempt is made to quantify and communicate the uncertainties as introduced into the subsurface representation by the experts themselves.

2.2 Harmonization of geo-information

The information, which is collected throughout the lifecycle of an infrastructural project, is needed for the characterization of the (sub-) surface as well as the design of the construction. In any larger infrastructural project, it is of major importance to be able to integrate complex systems that are consisting of 1) existing natural and man-made formations and 2) engineering structures (traditionally designed in CAD) (Orlic, 1997; Oosterom et al., 2006). With it, accurate data integration and the harmonization of geo-information is essential in order to make the design of a proposed engineering structure economic and safe.

For an economic planning of an infrastructural project, easy communication as well as data and information exchange needs to be guaranteed. To achieve this, real world representations should be equipped with sufficient metadata ('data about data') describing their meaning and implications for the development of the project in a language understandable by all different parties. For that reason, the harmonization of various types of geo-information is an important matter within the geo-information community. Consistent application of terms is a prerequisite for successful implementation and unambiguous adoption of legislation, regulations, guidelines and interpretations. A glossary shall therefore be established to define the meaning of those terms regarding geographic information that are used regularly within infrastructural projects. This will ensure that such terms are consistently and correctly interpreted, as far as is practicable, at all stages of the lifecycle of an infrastructural project.

A first step will be made with this research, which is embedded in the Dutch GIMCIW ('Geo-Information Management for Civil-Engineering Infrastructure') project. Additionally, strong connections can be established between this part of the research and parts of the INSPIRE initiative, which is being prepared inside the European Union since 2004 and will assumingly be started in 2007. The INSPIRE initiative for the development of an infrastructure for spatial information in Europe is also partly focused on the development and implementation of metadata as well as the development of a conceptual model containing harmonized geo-information.

This goal of geo-information harmonization becomes even more difficult to reach when considering the fact that the existing environment consists of two main components; i.e. surface and subsurface. The approaches to measure, represent, store and analyse surface or subsurface data are specific and quite different. In most cases, for the construction of engineering structures both representations have to be taken into account (see also Figure 2.1 and Figure 2.2).

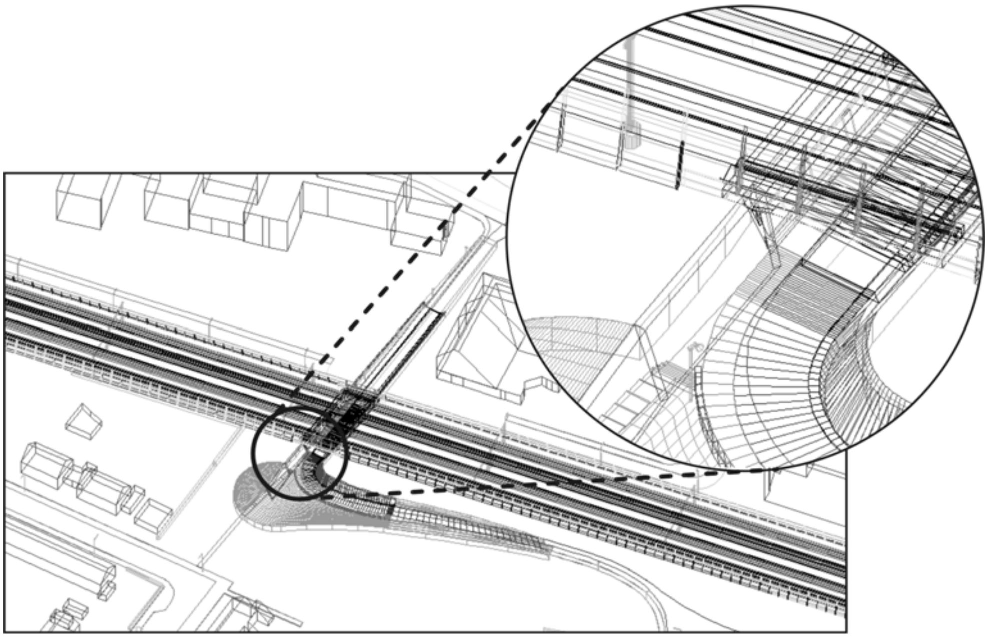


Figure 2.1 Construction design representation (CAD) for a cycle tunnel in Houten, The Netherlands (by courtesy of Holland Railconsult).

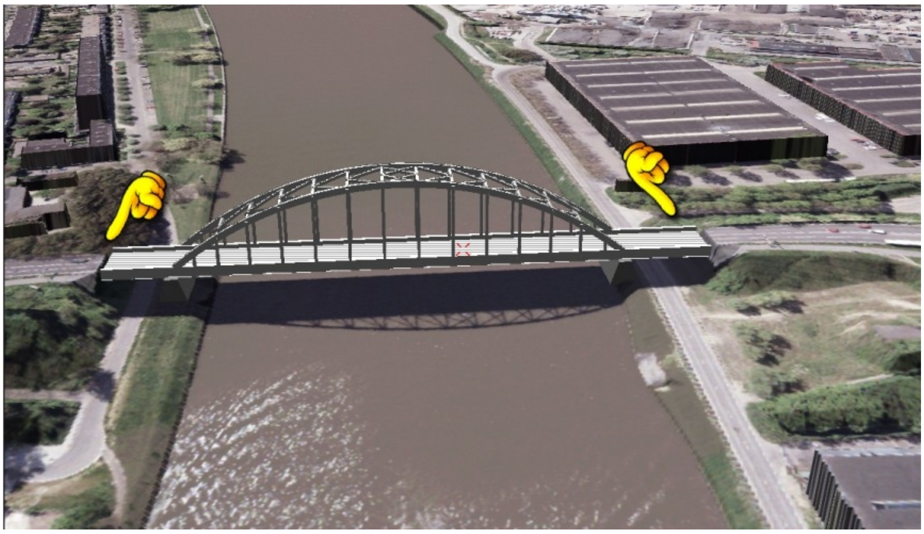


Figure 2.2 Bridge Amsterdam-Rijnkanaal; near Utrecht, The Netherlands (by courtesy of RWS-AGI).

In the 'underground world', emphasis is put on the best representation of the different geological formations and related processes that are often not available for direct measurement. Many complex geological conditions must be accounted for. This generally requires the application of numerical techniques, i.e. using three-dimensional interpretations of the underground that characterize the ground in terms of the spatial variability of physical, mechanical or other properties relevant to the problem treated (Orlic, 1997). When compared to 'subsurface objects', natural phenomena and man-made objects on the surface can easily be measured. The shape and size of surface objects is, in most of the cases, well visible. As a result, in the 'surface world' the emphasis is put on the best representation of the complexity of

the various objects; including information concerning their thematic properties and relationships.
One of the main problems in creating a harmonized system capable of offering a complete set of 3D functionalities is the amount of different data types. A (common) final representation should ideally contain knowledge about reality, so we have to consider the different types of real world objects it must represent (Raper, 1989).

2.3 Representation of real world objects

Regarding the diverse characteristics of different types of real world objects, we must consider applying different techniques for the representation of each object. What should be noted is that a spatial representation can be considered separate from the visualization of the spatial data, which is just a set of possible views over the representation (Jones C., 1989; Loudon, 1986); similar visualization techniques can also be applied to different representations in order to produce similar viewable results. Spatial queries and operations are performed against the different representations as required (Lattuada, 2006).
In the following sub-sections, the various representation techniques as applied for the representation of subsurface, surface and design objects will be presented. Several techniques might be overlapping and used for more than one object representation. However, the different representation techniques will still be presented for each of the objects in order to achieve a complete overview of the representation techniques as applied for the various objects.

Subsurface objects

Subsurface objects can have different dimensionality and can be represented by points, lines, surfaces and volumes (e.g. Orlic, 1997). Computer representations of subsurface objects have generally been subdivided in two classes that are namely: *surface (= boundary) representations* and *volume representations*.

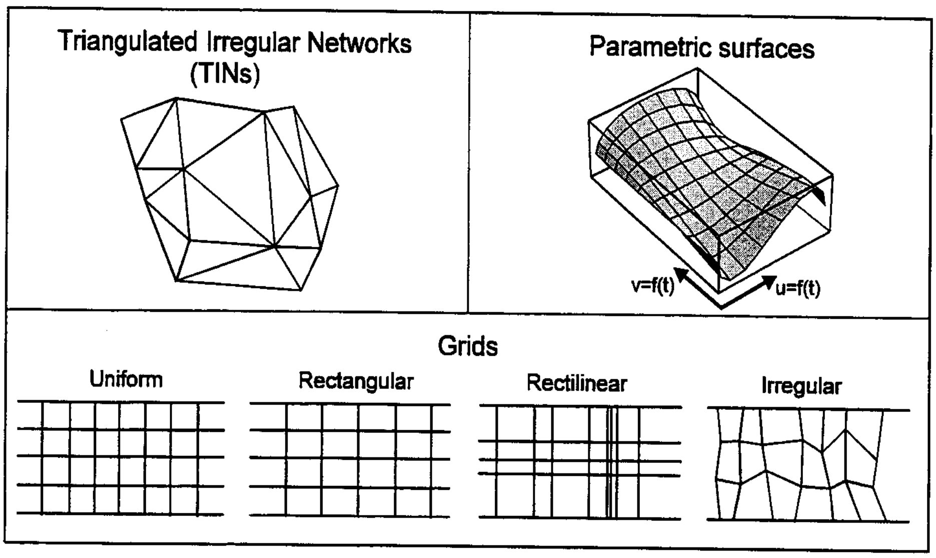


Figure 2.3 Surface representations (Orlic, 1997).

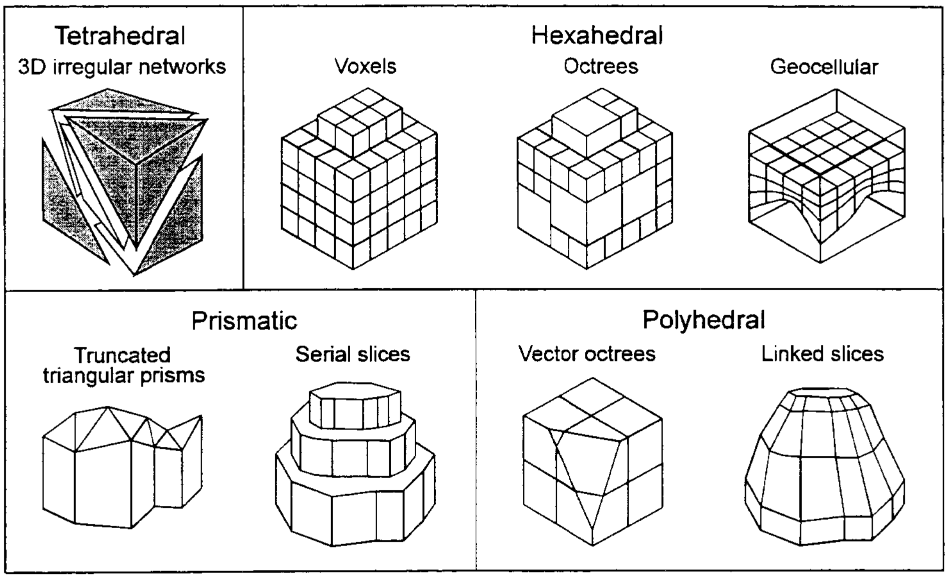


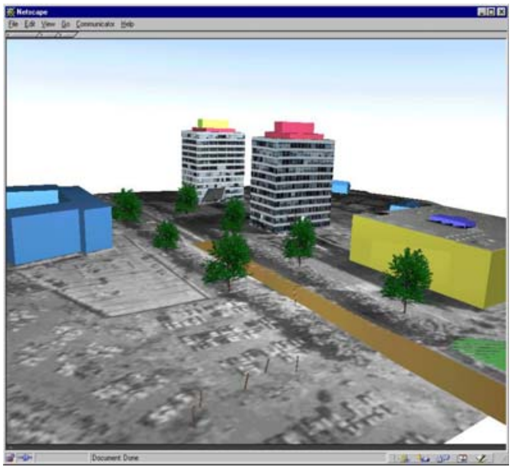
Figure 2.4 Volume representations (Orlic, 1997).

Surface representations (i.e. grids, parametric surfaces, TINs, etc.) are suitable for describing geometric characteristics of objects by surface entities (i.e. assuming the described volume is homogeneous) (e.g. Harbaugh and Merriam, 1968; Muller, 1988; Fried and Leonard, 1990), and volume representations (i.e. tetrahedral, hexahedral, prismatic, etc.) are suitable to characterize an object in terms of its internal properties, which can vary from one element to the next or from one element node to the next node (e.g. Requicha, 1980; Meier, 1986; Bak and Mill, 1989; Jones, 1989). Yet, subsurface objects often require features of both, surface and volume, representations, and most products today reflect these requirements by including elements of both types of representation techniques (e.g. Lattuada, 2006). With it, volume representations of a geo-object are commonly obtained by conversion from a surface representation (Orlic, 1997). Thereby, it is necessary that a surface representation can unambiguously (i.e. no gaps or overlaps in description of surface) describe the object. Generally, it is relatively easy to convert all volume representations and surface representations of subsurface objects, into volume representations, but mostly only approximately (Foley et al., 1990).
Looking at the standards available in the area of geology as well as geotechnology, several standards exist mainly concerning the classification techniques for rocks and also soils. Typical standards are for example the ISO 14689-1:2003 'Geotechnical identification and testing – Identification and classification of rock' or the ASTM D5878-05 'Standard Guide for Using Rock Mass Classification Systems for Engineering Purposes'. Frequently used standards are especially the British Standards, such as BS 5930 the 'Code of practice for site investigations' or the BS 1377 describing several laboratory testing methods.

Surface objects

Surface objects (i.e. objects on the surface) are usually represented in 2D and 2.5D with simple primitives, such as point, line and polygon (i.e. in vector representations) or as segmented areas (i.e. in raster representations). When considering 3D, real-world objects are described by four elementary objects (i.e. point, line, surface and volume). As suggested by the Open Geospatial Consortium 'OGC' (Open

Geospatial Consortium Inc., 2004), the simple primitives can be organized in *geometric* (simple features) or *topological* (complex features) data structures (see Figure 2.5).



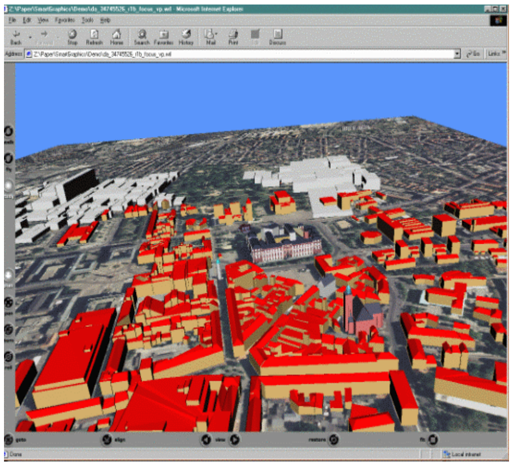


Figure 2.5 Two examples of the 3D topological models SSS and UDM.

The topological data structures (i.e. Formal Data Structure FDS, Tetrahedral Network TEN, etc.), as reported in the literature (e.g. Molenaar, 1989; Molenaar, 1990; Pilouk, 1996), can be subdivided in two main groups; that are, structures maintaining objects (object-oriented; i.e. relationships between objects must be derived) and structures maintaining relationships (topology-oriented; i.e. representation of objects must be derived). Frameworks (e.g. 9-intersection model, Egenhofer and Herring, 1990) to formally describe relationships between different objects (independently of the data structure) are also available (see also van Oosterom et al., 1994; Zlatanova et al., 2004).
Resulting from the work of several initiatives (i.e. such as the OGC), various standards have been made available to be used for the representation of surface objects and their attribute data (e.g. ISO, NEN, OGC, etc.). Especially the ISO[4] standards, international standards among others developed for the management of geo-information and the representation of surface objects, are frequently applied and can also be related to this research. Examples are:

- ISO 19107:2003 Geographic Information – Spatial schema
- ISO 19108:2003 Geographic Information – Temporal schema
- ISO 19113:2002 Geographic Information – Quality Principles
- ISO 19114:2003 Geographic Information – Quality Evaluation Procedures
- ISO 19115:2003 Geographic Information – Metadata
- ISO 19123:2005 Geographic Information – Schema for coverage geometry and functions

In the same way as the ISO, the OGC and NEN[5] have developed standards for the measurement and representation of surface objects as well as for the definition of proper geo-information exchange. The most relevant NEN standard for this research is probably the NEN 3610, i.e. the 'Basic scheme for geo-information – Terms, definitions, relations and general rules for the interchange of spatial information of spatial objects related to the earth's surface'. The Open Geospatial Consortium (OGC), for example, has, among other things, developed the document OGC 05-

[4] ISO – 'International Organization for Standardization'
[5] NEN – 'Nederlands Normalisatie-instituut'

087r4 'Observations and Measurements'. This document, which is, however, not yet a standard is dealing with a framework for measurements and encoding and is, therefore, of interest for the data collection of data, which will later be used for the development of real world representations of objects on the earth's surface.

Design objects

Design objects can be represented by various methods. A number of representational forms for three-dimensional objects have been developed in computer aided design. Some of these arise from applications and the data structure is determined wholly by the representational strategy. For example, in solid modelling, a method known as 'Constructive Solid Geometry (CSG)' is both a representational form and a method that facilitates a form of graphical interaction that enables engineering parts to be built up. The factors that the representation generally tends to determine are 1) the data structure (i.e. the form of the processing algorithm and the design of fixed program hardware), 2) the cost of processing an object, 3) the final appearance of an object, and 4) the ease of editing the shape of an object. In computer graphics, the most popular method for representing an object is the polygon mesh (triangular) representation (i.e. objects are approximated by a net of planar polygonal facets). In addition, the 'bicubic parametric patches', that are freeform curves and surfaces (i.e. objects are represented exactly by nets of elements called patches), 'constructive solid geometry' (i.e. objects are represented exactly by a collection of elementary parametrical objects, such as spheres or boxes), and 'space subdivision' (i.e. objects are embedded in a space where points in the space are labelled according to object occupancy) techniques are frequently used for the representation of design objects. With it, polygon mesh and space subdivision are approximate representational forms, whereas the other two techniques are exact representational forms. Furthermore, with the polygon mesh and bicubic parametric patch methods boundary representations are derived, whereas the other two techniques deliver volume representations (e.g. Watts, 1993).
Clearly, harmonization of the various types of representations is of great importance in a particular project (i.e. using comparable data structures and definitions), and would significantly facilitate the work of the engineers.

2.4 Software packages for real world representations

Since the numerical techniques as used for the development of the various three-dimensional representations are only feasible if performed by computer, they should ideally be created and available in digital form. Different software packages have evolved to meet the requirements of specific disciplines (e.g. GIS technology, engineering geology, civil engineering, etc.) and as such they provide various functionalities.

Geotechnical software packages

For the representation of real world objects in the subsurface, various software packages exist. In addition to the commonly used GIS packages (e.g. ArcGIS, ArcView, etc.), GIS and GIS-like packages (e.g. Lynx, Rockworks, GOCAD, etc.), which are specifically developed for the applications in the field of geology and engineering geology, are frequently used in order to achieve reasonable representations of the real world situation in the subsurface. These software packages are not only used to produce two-dimensional representations such as maps or cross-

sections, but also and even more increasingly for representations of the subsurface in three dimensions. Also, an increasing number of calculations and analyses are possible with the help of these software packages (e.g. basic statistics, layer thickness computations, etc.); however, most of them are still restricted to two- and 2.5-dimensions.
GOCAD, for example, is one of the geotechnical software packages frequently used for the representation of the subsurface. In general, it can be described as a program for the interactive construction and visualization of geological interpretations using point-, line- and area- data. The research concerning this software package has been started during the eighties when it became clear that traditional automatic mapping systems would never be able to represent complex surfaces and, more generally, complex geological volumes. At the same time, also experience using traditional CAD software developed for the car industry brought out its inability to accommodate the complex data encountered in geosciences. Following, the GOCAD software has been developed as an alternative to traditional CAD systems within the framework of the GOCAD research project.

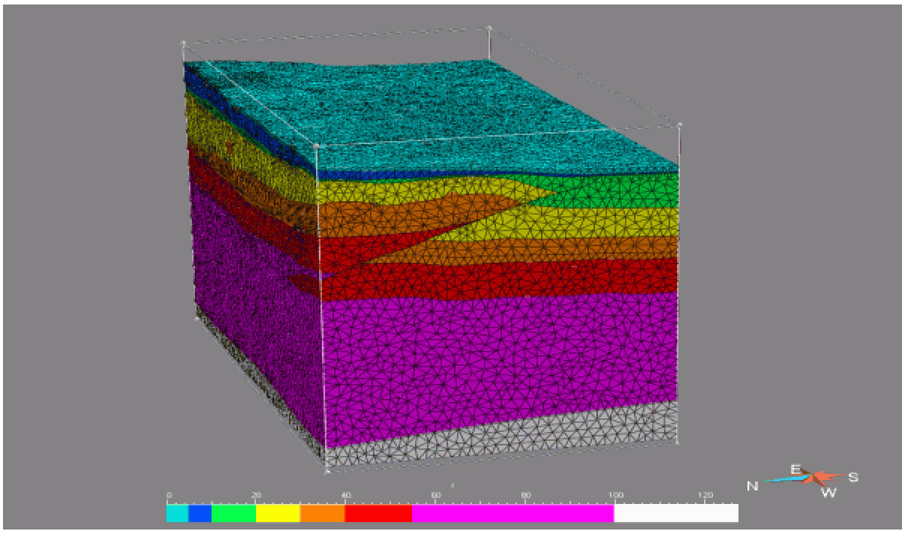


Figure 2.6 Subsurface real world representation in GOCAD (http://www.gocad.org/www/research/popup.php?photo=images_gallery/GocadResearch4.gif).

With the GOCAD program, a completely different strategy involving the discrete modelling of natural objects was proposed. In this discrete approach:

- The geometry of any object is defined by a finite set of nodes (= points) in 3D space
- The geometry of any object is defined by links bridging these nodes
- The physical properties of any object are defined as values and attached to the respective node; and not to volumes (e.g. in case of tetrahedrons) or surfaces (e.g. in case of triangles)

For example, if the object to be represented is composed of surfaces, then the links can be arranged in such a way that the mesh so defined generates triangular facets. These facets can be interpolated locally by flat triangles or, if needed, by curvilinear triangles. This strategy, apparently, can also be extended to the representation of curves and volumes. In practice, such a discrete approach is of no interest without a powerful mathematical tool able to interpolate the physical properties and the location (x ,y ,z) of the nodes defining the objects in the 3D space. Therefore, the rather robust interpolation- and approximation- algorithm DSI (= discrete smooth interpolation) has been developed, which allows the flexible generation and modification of 3D geometry objects in the subsurface. This interpolation method was specifically designed for the representation of natural objects, while taking into account a wide range of complex and more or less precise data.

Geographic Information Systems (GIS)

GIS systems have originally been developed for the (re-) construction, management and analysis of existing (geographical) objects (Lattuada, 2006), about which only sparse and incomplete information is available (i.e. real-world objects). GIS systems should be able to maintain information about spatial phenomena and provide means to analyse it and thus, gain knowledge of the surrounding world. In general, the tasks or the functions of a GIS are specified as follows (Raper and Maguire, 1992): 1) data capture, 2) data structuring, 3) data manipulation, 4) data analysis, 5) data presentation & distribution. For a long time, these functions were focussed on 2D. Nowadays, 3D GIS aim at the same functionality as 2D GIS, but in 3D space. Therefore, important developments in the field of 3D GIS are the improvement of 3D data collection techniques as well as developments concerning hardware, such as processors or memory and disk space devices, which have become more efficient in processing large data sets. Furthermore, elaborated tools to display and interact with 3D data are evolving. GIS software-tools have also made a significant movement towards 3D GIS (Zlatanova et al., 2002) and the major progress achieved in 3D GIS is on improving 3D visualization and animation (see also Figure 2.7).

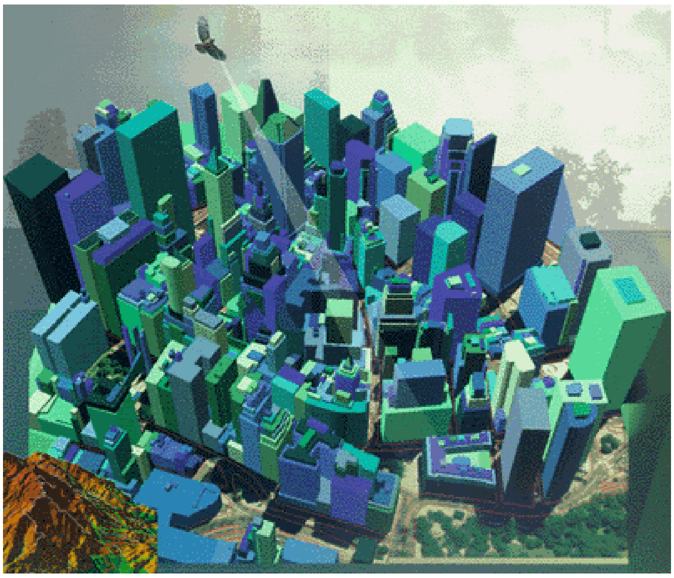


Figure 2.7 Urban development representation in GIS (http://www.dataplus.ru/soft/ESRI/ARCVIEW/3DAnalis.htm).

Nevertheless, 3D functionality, such as generating and handling (querying) 3D objects, 3D structuring or 3D manipulation and analyses, is still limited (or sometimes even lacking).

Computer Aided Design (CAD)

CAD systems, when compared to GIS packages, have been developed for the design of new, but well-defined objects (i.e. design objects and engineering constructions). In CAD oriented systems, the main target is to design objects using geometric primitives to design, evaluate, edit, and construct; classical CAD methods are used to design interactively curves and surfaces (Lattuada, 2006). CAD is a typical computer graphics tool for 3D design, which is used, for example, for car, machinery, the construction industry, and architecture (see also Figure 2.8).

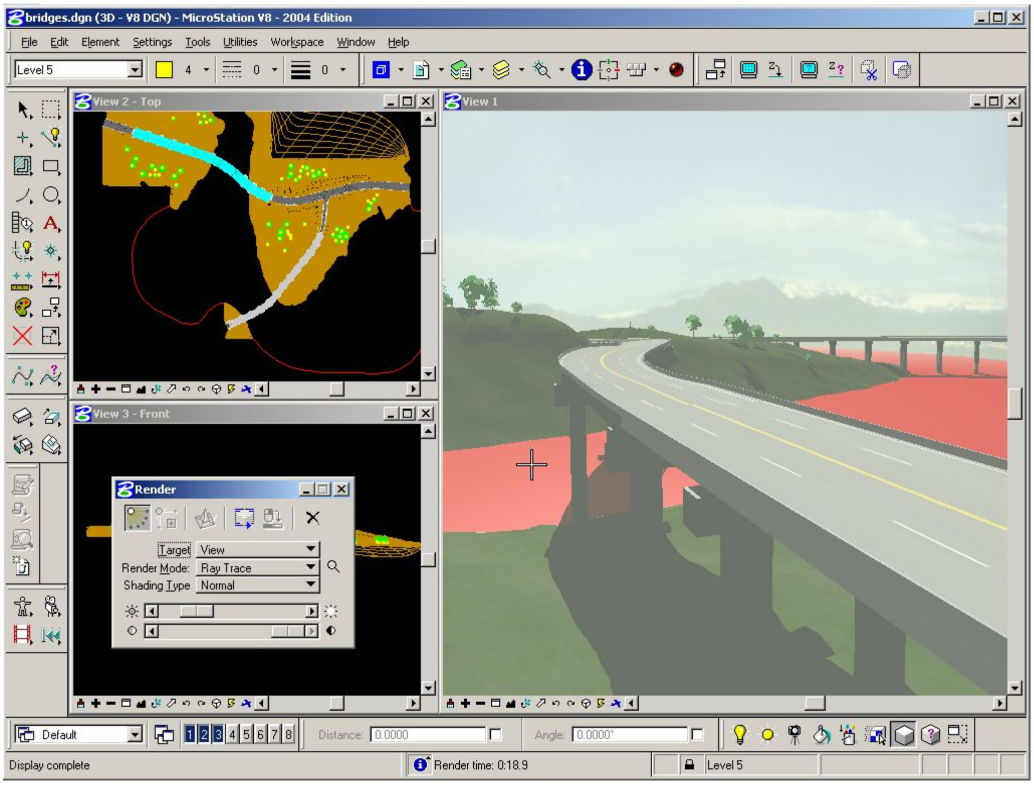


Figure 2.8 3D Bridge Design in MicroStation V8 (http://www.bentley.com/bentleywebsite/files/corporate/ydb/MicroStation-II.jpg).

CAD focuses on the geometric aspect of the object and its 3D visualization. The question arises whether CAD can support all the tasks required in the disciplines of geosciences (e.g. Corbea Diaz, 1996; Pilouk, 1996), as in CAD systems, the geometric description of objects is the first priority; while non-spatial information, especially topologic (i.e. topology in-between neighbouring objects) and attribute information, is not considered. Yet, this topology and attribute information supports many GIS operations such as query, search, or spatial analysis. As a result, the requirements for 3D data structures concerning CAD and GIS systems were different in the past; namely, GIS also required the storage of geometry and attribute data in one object, whereas this integration between these different types of information used to be less dominant in CAD. However, it should be noted that CAD systems for engineering applications are currently changing and, nowadays, tend to be called AEC systems (i.e. architecture, engineering and construction). With it, it should also be noted that many CAD systems (e.g. Bentley, AutoCAD) already offer more and more extended tools to represent and organize real world-objects. Vendors dealing with either spatial or thematic information attempt to achieve GIS functionality already for years and CAD vendors (e.g. AutoDesk, Bentley), nowadays, provide means to combine 2D and even 3D spatial data to thematic information and organize topologically structured layers.

As has been described above, the general understanding for GIS and CAD systems is changing. Besides analyses, GIS is becoming an integration of strong database management (i.e. ensuring data consistency and user control) and powerful editing and visualization environments (i.e. inheriting advanced computer graphics achievements); and also CAD vendors are trying to provide means to combine spatial as well as thematic data and organize topologically structured layers. Resulting from these developments in the field of spatial data management is, according to Zlatanova et al. (2002), also a changed role of Database Management Systems (DBMSs).

Database Management Systems (DBMSs)

DBMSs can generally be described as a computer program (or more typically, a suite of them) designed to manage a database (i.e. a large set of structured data) and to run operations on any collection of compatible, and (in case of a relational DBMS) ideally normalized, data of a particular application or problem (i.e. as requested by the numerous clients). According to Date (1995), the functions a DBMS should offer are 1) data definition, 2) data manipulation, 3) data security and integrity, 4) data recovery, and 5) data dictionary. At present, the so-called Geo-DBMSs can manage temporal, thematic *and* spatial data and they are providing spatial *data types* and *functions* that define the *spatial functionality* of this Geo-DBMS. A Geo-DBMS knows simple and composed data types, unfortunately, most of the spatial data types are still only two-dimensional; i.e. point, line, and polygon (Breunig and Zlatanova, 2006). According to the Abstract Specifications of the OpenGIS Consortium (OGC, 2006), the spatial object can generally be represented by two structures in the DBMS; that are namely, geometrical structure, i.e. simple feature, and topological structure, i.e. complex feature. While the geometrical structure provides direct access to the coordinates of individual objects, the topological structure encapsulates information about their spatial relationships. Each spatial object is then maintained in the DBMS environment following the Implementation Specifications of the OpenGIS Consortium (e.g. OGC, 1999).

Nowadays, DBMSs can be integrated within GIS systems and important parts of data analyses are performed by the DBMS during the execution of database queries. In contrast to GIS systems, in the field of CAD, the use of DBMS functionality was considerably more restrictive (Breunig and Zlatanova, 2006). However, an increasing number of CAD systems (e.g. MicroStation, GeoGraphics) have, at present, also developed extensions that make use of spatial data structures and functionality provided by Geo-DBMS. In general, it is to be expected that the significance of Geo-DBMS in GIS, CAD and geotechnical worlds will continue to increase, as there are a lot of reasons to use Geo-DBMSs in these systems (e.g. Breunig and Zlatanova, 2006). DBMSs can also be used as a bridge between GIS and CAD applications. But, in order to be able to provide a stronger management of objects from CAD and GIS, Geo-DBMSs have to extend their spatial support to accommodate design objects and real-world objects.

However, the question 'who is responsible for the spatial analysis' (i.e. front-end applications or spatial DBMS) is still open and even extensively discussed. Generally, it can be said that GIS functionalities that are not specific to a certain application belong in the DBMS and not in GIS (or CAD) front-ends (Zlatanova and Stoter, 2006). Nevertheless, it should not be forgotten that Geo-DBMSs, in the first place, are DBMSs; i.e. the location for data storage and management. Thus, the 3D functionality should not be completely taken away from GIS and CAD applications. 3D Geo-DBMSs should only provide the basic (simple) 3D functions (e.g. as computing volumes and finding neighbours as basis needed for querying) and application specific (i.e. as needed for complex analysis) should still be attributed to the GIS/CAD systems (Breunig and Zlatanova, 2006).

Concluding, it can be said that spatial objects of reality, nowadays, can well be visualized in the form of graphics in two- as well as three-dimensional representations. Unfortunately, the developments concerning data analyses were, up till now, less successful and they are still limited to 2D and 2.5D. However, as we live in a three-dimensional (3D) world, 2 as well as 2.5D representations and interpretations of project-related data are often not sufficient in evaluating construction sites for proposed engineering structures. Yet, none of the software packages that are commonly used can fulfil *all* requirements of earth science applications. The representation of complex systems that goes beyond drawing part of the problem well into the simulation, budgeting, environmental impact analysis and decision support makes a strong case for a tighter integration of the different software packages in a full three-dimensional environment (e.g. Lattuada, 2006). The various systems may well have features in common (e.g. all are concerned with geometry); however, they also differ in many respects such as dimensions, storage, analysis, semantics, etc. Accordingly, the combined use of the software packages often results in problems caused by non-compatible data structures or differences in scale and functional level.
Regarding the techniques available at present for the determination of data and representation qualities as well as the state-of-the-art of (sub-) surface and design representation techniques, this PhD research aims at the development of a suitable method for the quantification and communication of interpretation uncertainties in geotechnical real world representations. Furthermore, an attempt will be made to harmonize the various types of metadata and semantics as used in the different categories of representations (e.g. GIS, CAD, multi-dimensional representations, etc.). This approach should significantly facilitate the communication between the different parties working on the infrastructural project.

3 Research Question

The main research question as proposed for this study is:

- How can (meta-) data related to (sub-) surface and design objects be harmonized in the heterogeneous lifecycle environment of infrastructural projects (i.e. geology, geotechnology, surveying, remote sensing, civil engineering, etc.)?

The general research topic can further be subdivided into two areas of main interest that are namely:

1. Definition of metadata (e.g. including estimates of the interpretation uncertainty, etc.)
2. Harmonization of geo-information – with focus on metadata and thematic semantics

On the basis of this sub-division, the following research questions have been determined:

1. How can an efficient quantification and communication of the so-called interpretation uncertainties in subsurface geotechnical representations be achieved?
2. What kind of data structure is suitable for the management of the various types of geo-information and how can the different kinds of metadata be harmonized within this structure?

Definition of metadata

Specific research sub-questions that will have to be answered regarding the definition of metadata include:

- How can the 'interpretation uncertainty' of geological/geotechnical representations be determined?
- How can the different types of data and interpretation uncertainties be described and defined?
  - Regarding subsurface data (e.g. GIS, Lynx)
  - Regarding surface data (e.g. GIS)
  - Regarding construction design data (e.g. CAD)
- How can the quality of different types of representations be determined and communicated?
  - For subsurface representations (e.g. geology)
  - For surface representations (e.g. surveying)
  - For construction design (e.g. civil engineering)

Harmonization of geo-information

Specific research sub-questions that will have to be answered regarding the harmonization of geo-information include:

- How can varying data themes be included in one framework (harmonized model)?
- How can estimates of the interpretation uncertainties of the subsurface geotechnical representations be included in the metadata of these representations?
- How can uncertainty estimates regarding data and interpretations be included in the metadata of each representation?
- How can uncertainty estimates of the various representations be included in the metadata of each representation and equipped with additional semantic information in order to improve the communication between the parties involved in a project?
- Which software and data structure (i.e. objects, attributes, relations) is suitable for the harmonization of geo-information?

3.1 Hypothesis

Generated from these research questions are the following hypotheses, which will be tested in the course of this proposed study:

Definition of metadata

- Spatial data quality in subsurface, surface, and design data sets can best be described using existing methods (e.g. geostatistics, probability functions, fuzzy sets, etc.).
- Interpretation uncertainty is largely dependent on data quality and quantity and influenced by the experience and a priori knowledge of the interpreter. Furthermore, it is dependent on the local geological complexity and also on the scale of the project.
- Uncertainties of the different interpretations are dependent on the quality and quantity of available data as well as on scale and complexity of the project they are intended for.
- The uncertainty of different types of representations is best defined in the form of 'quality parameters' (e.g. as percentages).
- Expert knowledge and experience of various professionals facilitates interpretation of collected data and is best included as additional information in the real world representation.

Harmonization of geo-information

- Interpretation uncertainties are best defined in the metadata of the geological/geotechnical representations they have been determined for.
- The uncertainty of the various data and interpretations can best be attached in the metadata of the different representations.

- Semantic information attached to uncertainty and quality estimates of the different representations facilitates the communication between the different parties.
- Various data types are best integrated in an object-relational DBMS as part of the overall system architecture.
- The different representations (i.e. subsurface, surface, design) still cannot be combined in one integrated system; however, all data and information concerned by a certain project can be managed with suitable uncertainty indications in one common database management system. This facilitates the communication between the different parties involved in the infrastructural project.
- An agreement of models for different themes can be achieved.

3.2 Relevance

The inefficient use of information in large civil engineering infrastructural projects causes an increase in working hours and, subsequently, project costs. With a harmonized geo-information structure, including expressions concerning the uncertainties in foremost subsurface real world representations, the exchange of information can significantly be facilitated and, hence, an easier cooperation of the different parties involved in these projects is possible. This PhD research is undertaken in order to develop a suitable method for the quantification and communication of interpretation uncertainties in geotechnical real world representations. Furthermore, an attempt will be made to harmonize the various types of metadata and semantics as used in the different categories of representations (e.g. GIS, CAD, multi-dimensional representations, etc.). This approach should significantly facilitate the communication between the different parties working on the infrastructural project.
As has already been mentioned before, this work, which is embedded in the RGI project 'Geo-Information Management for Civil-Engineering Infrastructure (GIMCIW)' on national level, is in close relation to the approach of the section GIS-technology at the research institute OTB of the TU Delft to develop the technology to stimulate the realization of the geo-information infrastructure (GII). On an international level, this section goal is in one line with the goal of several initiatives and projects focusing on the development of the so-called 'Spatial Data Infrastructure (SDI)'.
In more general terms, this research project can be connected to various other projects in- and outside the Netherlands. Inside the Netherlands, close relations can be found with various projects of the RGI initiative as well as parts of the PIM[6] program. On an international level, this research project can be connected to, among others, the INSPIRE initiative and the HUMBOLDT project.

[6] PIM – 'Partnerprogramma Infrastructuur Management'

4 Research Methodology

This research is part of the BSIK project 'Geo-Information Management for Civil-Engineering Infrastructure (GIMCIW)'. Various (engineering) companies are involved in this project, several, for example, by contributing data from specific construction locations. This data will be used throughout this study in form of different case studies. Possible case studies to be considered are:

1. Betuweroute
2. Groene Hart Locaties
3. Nieuwe Woonwijken

In general, these are all large infrastructural projects located in the Netherlands. But, up till now, a final decision on the use of one specific case study has not yet been made. This will be done in the near future during a meeting with the consortium members that are concerned by this decision.
Besides the implementation of this research project in the GIMCIW project as part of the Dutch RGI initiative, a strong connection can be found between this work and the European INSPIRE initiative, which is also concerned with the establishment of an infrastructure for spatial information; but then on an international (European) level.
This research is based on the topic of data harmonization in large civil engineering projects (see also chapter 3). In addition, the study can be subdivided into two main components that are namely: 1) Meta data (e.g. quality, accuracy, uncertainty) and 2) Harmonization of Geo-Information with focus on the meaning of data (i.e. thematic semantics of data). Work on this research will be started with phase 0; however, this sub-division should not be taken as a timeline. Work regarding the two different phases might as well be overlapping in case needed.

4.1 Phase 0: Evolving an overview of the different specializations

During the first phase of the research, an overview of the various specializations involved in infrastructural projects will be approached. With it, it is important to collect information about all

- data types/themes (collection of the various object and attribute types)
  - object methods/functions (e.g. for data retrieval, etc.)
  - object relationships
  - object constraints
- data structures
- data exchange formats
- representations

that are used by the various parties for the development of the infrastructural project and hence, that are included in the lifecycle of the specific infrastructural project that is of interest for this study. In this phase of the project, it is important to determine possible relationships and/or overlaps between the various data sets. Shared object definitions might later facilitate the harmonization of all geo-information. Furthermore, it should be determined if and how metadata is included in the various data sets and how this metadata can then be used in order to attach uncertainty information to the different real world representations.

4.2 Phase 1: Metadata

In this phase of the research, uncertainties regarding the real world representations (e.g. GIS, CAD, multi-dimensional representations, etc.) developed throughout an infrastructural project will be determined. Focus will thereby be put on the determination of interpretation uncertainties within subsurface real world representations.

Similar to this project, parts of the INSPIRE initiative are focused on the harmonization of geo-information and, additionally, the determination of quality and uncertainty information concerning subsequent representations. Foremost the 'Drafting Team 1' of the INSPIRE initiative keeps itself busy with the development and implementation of metadata. The goal of this team is 'the technical content for the Metadata Implementing Rules'; with focus on the 'semantic description of the metadata elements to describe spatial resources'. The Implementing Rules apply to spatial datasets, datasets and series. In addition, they are applicable to the functional concepts of discovery, evaluation and use, and define the minimum requirements for implemented metadata (DT Metadata, 2006).

Before starting the determination of uncertainties of different real world representations and data sets, a thorough literature study concerning uncertainty will be done. This has already been started with throughout the last months. However, the field of uncertainty is rather wide and the literature study still needs to be extended. As has been mentioned before, many different kinds of uncertainty (e.g. uncertainty to spatial prediction, uncertainty resulting from site investigations/ surveys/ measurements, uncertainty resulting from geological interpretation etc.) exist regarding the data collected for large infrastructural projects. With the help of the literature study, the different kinds of uncertainty should be classified and knowledge about the various methods (e.g. geostatistics, fuzzy set theory, probability theory, etc.) already existing for the determination of different kinds of uncertainties collected. Based on this literature study, the most suitable method for the definition of the interpretation uncertainty in geotechnical subsurface representations as well as the determination of possible uncertainties (i.e. due to a lack in spatial data quality, measurement inaccuracies, etc.) in surface representations will be established.

Then, uncertainties concerning geotechnical subsurface interpretations (with focus on the so-called interpretation uncertainty) will be determined. The determination of uncertainties in subsurface interpretations is assumed to be the most difficult and also most important task when compared to surface or design representations, since geotechnical (subsurface) representations are usually based on a limited number of data.

It should be realized though that throughout this research only a first step will be made towards the complete definition of interpretation uncertainties in subsurface representations. This in the same way as the determination of uncertainties regarding surface and design representations is out of the scope of this research as also another topic needs to be covered during this research (see also section 4.3). Therefore, the approach as followed in this research is to describe the level of interpretation uncertainty in a certain subsurface representation on a scale of, for example, 1 to 5; with 1 a low level of interpretation uncertainty and high reliability of the subsurface representation and 5 vice versa. To be able to get an estimation of the interpretation uncertainty and to derive an estimation of its influence on the quality of the subsurface representations, several steps will be followed. The different steps will be described in the following paragraphs.

Collecting experience and knowledge

To start with, knowledge will be obtained from a number of people that either develop or use geotechnical representations. In order to get a feeling for the influence that experience and a priori knowledge have on developed interpretations, structured interviews will be undertaken with various engineers as well as end users (i.e. mainly with partner organizations inside the GIMCIW project). Thereby, information about the influence of the human being (i.e. its knowledge and experience) on the interpretational process needs to be determined. In addition, insurance companies (i.e. large companies like AEGON) will be contacted in order to learn from the experiences they have made in, for example, risk analyses. Also, it needs to be determined what kind of factors these companies take into account that might lead to a higher uncertainty in a representation and, hence, influence the safety of a certain structure. In the course of the GIMCIW project, in which this PhD research is embedded, the results of the structured interviews will be summarized in a GIMCIW project report and as such distributed amongst the project partners. The results will be presented in a qualitative way rather than in quantitative expressions.

Comparing real world representations with situations in reality

During this phase of the research possibly two different case studies (i.e. infrastructural projects inside the Netherlands, in which GIMCIW project partners are involved) will be analyzed in order to acquire information about the influence of the expert knowledge on the quality of a real world representation. Hereby, data representations developed for the different case studies will be evaluated in comparison with the situation of the real world. During this step, it needs to be determined, how far the developed geotechnical subsurface interpretations (e.g. rock/soil layer boundaries, layer thickness) correspond to the situation that can be found in reality. With it, the quantity as well as quality of the data available for the interpretational process needs to be taken into account. It is expected, for example, that the less the quantity of the data set, which can be used for the interpretation process, the greater will be experience and knowledge the geology expert needs to apply and the greater might be the uncertainty he/she brings into the interpretation. Also, differences in scale (e.g. map scale; i.e. level of detail needed) will have influence on the interpretation uncertainty, as in small-scale representations; the expert's knowledge might have a larger influence than in large-scale representations where less detail needs to be included.

Determination of uncertainty expressions

After experience values have been derived from people involved in the different parts of an infrastructural project and after comparing developed representations with reality, 'uncertainty estimates' need to be determined. These values should describe the interpretation uncertainty that a geoscientist equipped with a certain a priori knowledge would bring into an interpretation when trying to achieve a real world representation of the subsurface according to the situation found in reality. This interpretation uncertainty of the subsurface real world representation needs to

be expressed in a suitable way. To achieve this, the level of interpretation uncertainty in a certain subsurface representation will be defined on a scale of, for example, 1 to 5; with 1 a low level of interpretation uncertainty and high reliability of the subsurface representation and 5 vice versa. The uncertainty information regarding the subsurface real world representation should be extended with additional information. This should give an indication on the meaning of the parameters and might later be extended with information concerning the interaction of the construction with the environment and possible hazards.
As has already been mentioned before, only a first step will be made in this way towards the complete definition of interpretation uncertainties in subsurface representations. The complete definition of uncertainties in subsurface as well as surface and design representations is out of the scope of this research as also other topics will be covered during this research (see also section 4.3). The approach towards the determination of uncertainties regarding subsurface representations, however, will be made, since the impact of interpretation uncertainties, as introduced into the representation by the geology expert, is still unknown. It is to be expected that interpretation uncertainties have a significant influence on the quality and thus on the (re-) usability of the representation.
When compared to subsurface data sets, uncertainties in surface data sets can be considered smaller and consequently also their influence on the quality of real world representations. That is mainly caused by the fact that in surface interpretations, the situation is less complicated as in most cases the objects, which are to be captured, can be seen on the surface and measured directly. Furthermore, a number of reliable interpretation rules are already available and the quality of the final representation is less dependent on the skills of the expert. Extensive research has already been undertaken in the field of uncertainties in surface representations; e.g. concerning spatial data quality in surface real world representations. For further information about this topic it can be referred, for example, to the recently finished PhD thesis by Pepijn van Oort (Oort, 2006) or the ongoing MSc research project by Erik de Ruiter in the course of the GIMA program. Oort (Oort, 2006), for example, covered two topics in his thesis; that are namely 1) the description of spatial data quality and 2) the understanding of the implications of spatial data quality.
Also, the definition of uncertainties included in design representations will not be covered. For a construction design no data from the real world needs to be collected or used and, hence, the uncertainty regarding these designs can be assumed zero. Since design objects are developed and designed by specific computer programs (e.g. CAD) without the use of any real world data, their quality must be high. The only case, in which the quality of any design is questionable, occurs when the construction designs need to be changed according to the conditions met at the construction site. In this case, the design representation needs to be adjusted; otherwise, the quality of the construction design is decreased.

4.3 Phase 2: Harmonization of geo-information

This phase of the research aims at the development of a harmonized system suitable for the combined management of all kinds of geo-information (i.e. concerning geology, geotechnology, GIS technology[7] and civil engineering) used throughout an infrastructural project, and the development of one harmonized 3D data description (i.e. including subsurface, surface and design representations). This would

[7] GIS technology includes, among other, elevation, topographic, ownership, etc. information

significantly facilitate the communication between the different parties in an infrastructural project. Moreover, it needs to be determined where in this management system uncertainty estimates regarding the individual real world representations can be kept. In this way, an intersection should be made between the various object types and descriptions and the metadata, such as uncertainty information, etc.
In general, this part of the research has close relations to the Dutch standard NEN 3610, in which a basic scheme for geo-information is defined. With it, terms, definitions, relations and general rules for the interchange of information of spatial objects related to the earth's surface are identified (see also NEN 3610, 2005). Also the European INSPIRE initiative is closely related to this part of the PhD research, especially considering the activities of the drafting team 'Data Specifications'. This drafting team is busy with the definition of a conceptual model specifying the generic aspects of geometry, topology, identifiers and relationships between spatial objects. This generic conceptual model should then be used as the reference point for further data specification and harmonization processes, and will provide a common baseline for the representation of the spatial, topological and temporal characteristics of the modeled real world entities.
The work in this phase can be subdivided in several smaller steps, which will now be further explained. Before starting the harmonization and implementation process, the software, which is intended to be used, needs to be studied. It is considered to be very important to be familiar with the software, since many decisions, which will have to be made throughout the study, require a strong knowledge of the different software packages (mainly DBMSs, such as OracleSpatial).

Integration of uncertainty estimates

In this step, it needs to be decided on where and also in which form the established uncertainty estimates of the different representations can be saved. At present, it is suggested to keep these uncertainty estimates as part of the metadata of each representation (e.g. subsurface, surface, and construction design). However, a final decision can only be made, when the software, which will be used, has been studied thoroughly and when its characteristics are known better.

Harmonization of different types of real world representations, attribute information and metadata

In order to find the most suitable way for the harmonization of geo-information as developed and used in infrastructural projects, several decisions will have to be made concerning the use of data types, metadata, different representation types, etc. With it, it is very important to decide on which is the best way to harmonize the different data types and representations; i.e. regarding the choice of data types, object representations (e.g. geometrical or topological in a DBMS), or the storage of attribute information (e.g. directly at the object or central at DBMS level) and metadata. Therefore, different types of attribute information as well as metadata as used in the various representations (i.e. subsurface, surface and design representations) and, hence, throughout the various disciplines (i.e. Geotechnology, GIS technology, Civil Engineering) will have to be compared in order to arrive at a good solution for the harmonization of geo-information.
Since a harmonization of the various object (i.e. geometry) as well as metadata (i.e. semantics) types is out of the scope of this work, it has been decided to focus this research especially on the harmonization of semantical data; meaning attribute

information attached to the different representations as well as additional metadata providing 'data about data'. The harmonization of the various geometry objects will be accomplished by another PhD student at the GIStechnology-section at the TU Delft.
The actions that will be taken during this research in order to partly achieve the harmonization of geo-information are the following:

Choice of appropriate software packages
For a successful harmonization of the various types of geo-information (e.g. GIS, CAD, multi-dimensional representations, etc.) that are developed and used throughout the lifetime of an infrastructural project, it is suggested to use a Database Management System (e.g. OracleSpatial). For any analysis on the data, a suitable GIS system (e.g. ArcGIS) will be used. Furthermore, design representations should preferably be developed in a common CAD system (e.g. MicroStation).
This decision is motivated by the fact that Geographical Information Systems (GISs), nowadays, are sophisticated systems for maintaining and analyzing spatial and thematic information on spatial objects. Moreover, DBMSs are increasingly important, since DBMSs are traditionally used to handle large volumes of data and to ensure the logical consistency and integrity of data, which also have become major requirements in GIS. Consequently, DBMS occupies a central place in the new generation GIS architecture.
For this study, it has been decided to use ArcGIS as GIS system, OracleSpatial as DBMS and MicroStation as CAD system. The main reason is that these are widely used systems offering the required functionality for this work. In this study, it is of major importance to use common systems in order to apply the developed methodologies on case studies from the industry. This is facilitated if the same computer systems are used. Furthermore, the above mentioned systems as well as software licenses are already available at the TU Delft, which saves time and money that would otherwise be needed to purchase these systems.

Choice of widely used attribute types
Since the harmonization of the various geometry objects will be covered by another research, this work is focused on the harmonization of attribute data as well as additional metadata (i.e. semantical data) concerning the various real world representations. With it, similar attribute types (and tables) should be used for the management of additional information concerning the captured objects and their relationships.

To achieve this goal of semantic data harmonization, an overview must be gained regarding the different types of attributes that are used in the various real world representations. Since significant differences exist between the objects that must be described within the disciplines and even more regarding the accessibility and quantity of data available for these representations, different approaches for the set-up and organization of attribute data can be assumed. Widely used attribute data types are, among others, characters (e.g. words and sentences), strings, numbers, colors, dates, Boolean (i.e. Binary values, such as true/false or 1/0). Another important aspect is the method applied in the various disciplines in order to create and save this attribute information. It needs to be determined, if the attribute information is kept directly with the object representation (e.g. within the CAD or GIS system) or saved central at a database level (e.g. using OracleSpatial).
To this end, a questionnaire will be distributed between the project members and information collected about their common approach to develop and save attribute information. After the information is gathered, a common way will be developed for

the description and saving of attribute data. As assumed now, the attribute information concerning single objects will be kept at an object level. Attribute information concerning real world representations, on the other hand, will be managed at a central database level, using OracleSpatial and the SQLDeveloper. Thus, a common framework (based on the same generic model) needs to be developed in order to achieve the harmonized way of geo-information handling.

Choice of appropriate metadata types

As has also been mentioned in the second chapter, different methods can be used in order to obtain representations of the subsurface, surface and construction design. For example, surface representations techniques (e.g. grids, parametric surfaces, TINs, etc.) or volume representations techniques (e.g. tetrahedral, hexahedral, prismatic, etc.) can be used to make real world representations of the geotechnical situation at the construction location.

In addition to a harmonized handling of attribute information, similar metadata types (and tables) should be used for the management of additional information concerning the different representations, subsequent attribute information and their subsequent meaning (e.g. regarding the influence of sub-/surface conditions on the safety of the construction and/or the whole project, etc.). Thereby, it needs to be established in how far the various parties make use of metadata to facilitate the communication of their data and representations. Furthermore, it needs to be determined, if, for the development and implementation of metadata, specific metadata schemes (e.g. Dublin Core or MODS) as well as syntax rules (e.g. SGML or XML) are applied. Based on this knowledge, a suitable way for the harmonized management of metadata in infrastructural projects needs to be defined.

Throughout this research, metadata will most probably be saved at a central database location, using OracleSpatial. Moreover, the ISO standard 19115 will be followed. In this standard, the mandatory and conditional metadata sections, metadata entities and elements are defined. In Figure 4.1, for example, the class 'MD_Metadata' is defined and relationships with the other metadata classes are shown which, in combination, define metadata for geographic data.

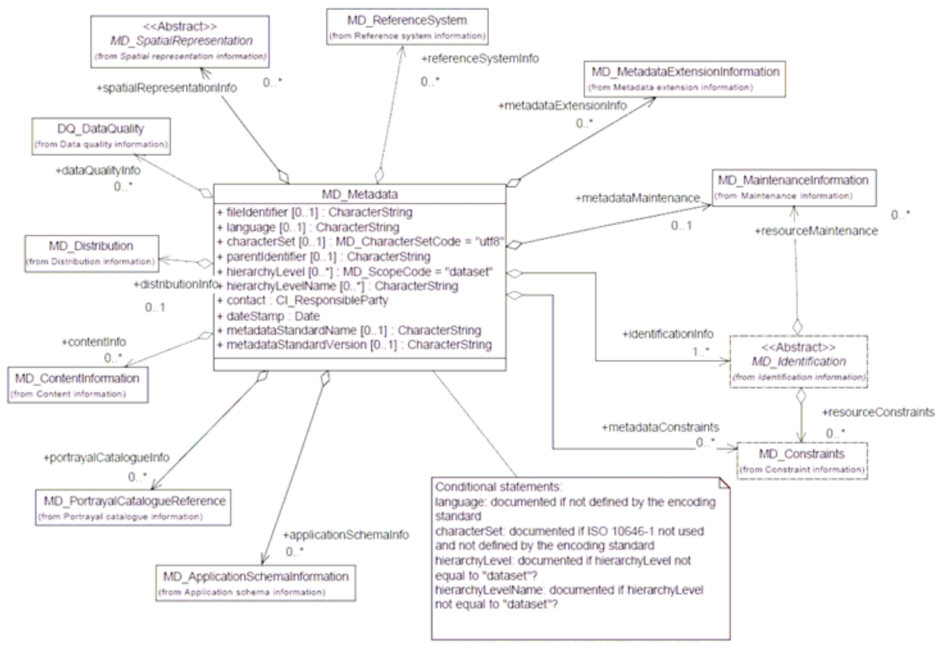


Figure 4.1 Metadata entity set information (ISO 19115:2003).

Implementation in the DBMS

As has been mentioned before, the different geo-information-types (e.g. concerning subsurface, surface and design representations) as used throughout the lifetime of an infrastructural project, should be harmonized in a common framework (based on the same generic model). In order to enable the harmonization of different geo-information, the GIS-technology, including sub-units, such as Geo-DBMS or GIS packages (e.g. editor, analytical tools, etc.), must be adjusted. Ideally, the behavior, characterization and interpretation of the surface, subsurface and construction design from the CAD world should be combined in complete 3D (or even 4D) representations.

A first step will be made with this research, where the harmonization of attribute information as well as metadata is approached. With it, all parties involved in a project manage their information in such a way that the data is always easily accessible and understandable. It is suggested that the data is kept at a database level. Additionally, meta-information should be available describing the all features of the data set as well as the indication it gives towards a certain project.

Furthermore, uncertainty information about the different real world representations is of great importance. Possible inaccuracies of defined properties in the subsurface interpretation and the influence of these possible deviations on the execution of the project must, for example, be determined in order to allow a risk analysis of the project. Especially the (interpretation) uncertainties regarding subsurface real world representations are to be determined during this work. These should be included in each representation with the help of metadata tables kept either at DBMS level or in the same table with the data. The uncertainty estimates in the metadata tables should be extended with semantic data. This should give an indication on the meaning of the estimation and might later be extended with information concerning the interaction of the construction with the environment and possible hazards.

Concluding it can be said that most of the data for the harmonized geo-information will be kept at DBMS level. The parties involved in the infrastructural project should implement their data in such a way that it will always be easily accessible and understandable by every user. Therefore, meta-information should be available at DBMS level describing the most important features of the various data sets as well as the indication it gives towards a certain project. Furthermore, uncertainty expressions concerning the different real world representations should be included in the meta-information, preferably extended with semantic data giving an indication on the meaning of the uncertainty estimation.

5 Proposed Output

5.1 Articles

It is expected that from the proposed study a number of reports, articles and also papers will be prepared for international journals and congresses that are namely:

- Regular reports inside the BSIK RGI project group
- 4 or 5 articles in (international) congresses
    - e.g. IAEG, ISPRS, etc.
- 2 papers in reviewed scientific journals
    - e.g. Computational Geosciences, Computers and Geosciences, Bulletin Engineering Geology and the Environment, etc.
- PhD thesis

5.2 Conferences

It is projected that during a period of approximately three years (07/2006-07/2009) one or two conferences per year will be attended. However, at present it cannot be foreseen which conferences these exactly will be. Certainly, the attendance at conferences should evenly be subdivided between the fields of 'Engineering Geology' and 'GIS Technology' in order to cover both aspects of this research.

In the field of Engineering Geology, several conferences of the 'International Association for Engineering Geology and the Environment (IAEG)' might be of interest for this study, such as the IAEG conference 'Engineering geology for tomorrow's cities' or the 'Fifth IAEG Asian Regional Conference on Engineering Geology for Major Infrastructure Development and Natural Hazards Mitigation'. Another interesting conference could be the IAMG[8] Annual Conference 'Quantitative Geology from Multiple Sources' and the '11th Congress of the International Society for Rock Mechanics (ISRM)'.

Regarding the field of GIS technology and the problem of data harmonization in this study, the GIScience conference should be considered as well as conferences concerning spatial data handling (SDH). Also the 'International Conference on Scientific and Statistical Database Management (SSDBM)' could be considered. Furthermore, it is considered to attend the '5th International Symposium of Spatial Data Quality' at ITC Enschede in June 2007.

5.3 Other types of results

Besides the expected output in the form of reports, articles and papers, various other types of results are to be expected from this research project, such as:

[8] IAMG – International Association for Mathematical Geology

- System prototypes:
  i.e. concerning possible software and program prototypes as developed throughout this research for the interpretation of geo-information as collected and used in infrastructural projects.
- Harmonized data descriptions:
  i.e. concerning a common method for the description and interpretation of the various types of geo-information as needed in infrastructural projects (i.e. subsurface, surface and design data), for example in the form of "UML" models.
- Data sets:
  i.e. concerning data sets that are saved in a harmonized environment and ideally equipped with uncertainty estimates.

6 Work Plan

Activity/ Year		2005	2006				2007				2008				2009		
		Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3
	Literature review																
	Report writing (RGI)																
	Proposal writing																
	Proposal defence																
	Article I																
	Learn Software[9]																
PHASE 1	Literature Uncertainty																
	Interpretation Uncertainty[10]																
	Article II																
	Overall Uncertainty[11]																
	Article III																
	Paper I																
PHASE 2	Implement Uncertainty[12]																
	Article IV																
	Data Harmonization[13]																
	Article V																
	Paper II																
	PhD thesis																
	PhD Draft																
	PhD Defence																

Suggested titles of the various papers and articles are:

[9] Particularly, introduction to Oracle Spatial and MicroStation GeoGraphics.
[10] During this part of phase 1, a suitable method needs to be developed to quantify the so-called interpretation uncertainty in subsurface representations.
[11] During this part of phase 1, other types of uncertainties in real world representations, such as spatial data quality, need to be determined and, in combination with the interpretation uncertainty, quality parameters for (sub-) surface and design representations established.
[12] During this part of phase 2, a suitable method needs to be determined for the implementation of the various uncertainty estimates and quality parameters into the respective data sets and real world representations.
[13] During this part of phase 2, a suitable method needs to be determined for the harmonization of the various types of geo-information as collected and used in infrastructural projects.

- Article I: "Spatial Information technologies for modeling the real world (including the subsurface) and the use of integrated DBMSs for an efficient data management"
- Article II: " A method for the quantification of interpretation uncertainties in subsurface representations"
- Article III: "A method for the determination of quality parameters of subsurface representations"
- Paper I: "Determining quality parameters for real world representations of surface, subsurface and construction design in infrastructural projects."
- Article IV: "Implementing uncertainty estimates and quality parameters into data sets and real world representations of infrastructural projects."
- Article V: "Harmonization of geo-information in infrastructural projects."
- Paper II: "Harmonization of geo-information in large infrastructural projects – with focus on the quality aspect of geo-information."

7 Budget Plan

7.1 PhD Research Budget (ITC)

<table>
<tr><th>Budget's draft (Euro)</th><th>2005</th><th>2006</th><th>2007</th><th>2008</th><th>2009</th></tr>
<tr><td>Laptop + accessories</td><td colspan="5">3000</td></tr>
<tr><td>Office materials</td><td colspan="5">500</td></tr>
<tr><td>Housing</td><td colspan="5">9000</td></tr>
<tr><td>Travelling (Enschede-Delft)</td><td colspan="5">2500</td></tr>
<tr><td>Total</td><td></td><td></td><td></td><td></td><td>15000</td></tr>
</table>

7.2 Other costs

Expenses (Euro)	Responsibility
Salary	Completely transferred by ITC - agreement between ITC & TU Delft; salary shared; 50 % paid by ITC, 50 % paid by TU)
Housing	Enschede: Permanent; paid by PhD Delft: Paid from ITC Research Budget (see above)
Office Space	Enschede: Paid by ITC Delft: Paid by TU Delft
Supervision	Out of my responsibilities
Software (if necessary)	Provided by ITC and TU Delft
Data sets	Provided by co-operating companies in GIMCIW project
Academic events	Paid from project budget (GIMCIW)
Co-operation/ Visits (e.g. Sweden, USA)	Paid from project budget (GIMCIW)

8 Project Organization

8.1 General Organization

As has already been mentioned, this research is part of the BSIK project 'Geo-Information Management for Civil-Engineering Infrastructure (GIMCIW)', which is one of many projects in line of the RGI ('Ruimte voor Geo-Informatie') concept in the Netherlands. Various (engineering) companies (e.g. TNO-NITG, DHV, Fugro, etc.) are involved in this project, several, for example, by contributing data from specific construction locations. This data will be used throughout this study in form of different case studies.
In general, these are all large infrastructural projects located in the Netherlands. But, up till now, a final decision on the use of one specific case study has not yet been made. This will be done in the near future during a meeting with the consortium members that are concerned by this decision.

8.2 PhD Research Organization

The PhD research is undertaken in co-operation between ITC Enschede and OTB at the TU Delft. Workspaces for the PhD student are provided at both locations and the PhD student spends approximately 50% of its time in Enschede and the other 50% in Delft.
The university promotor is *Prof. Dr. Ir. Peter van Oosterom*, from the Onderzoeksinstituut OTB (Section GIS technology) at the TU Delft, and the ITC promoter is *Dr. Robert Hack* Head of the Geological Engineering specialization.
Furthermore, *Dr. Sisi Zlatanova* is taking care of the daily supervision at the Onderzoeksinstituut OTB (Section GIS technology) at the TU Delft.

9 References

Al-Abdalla, M. (1998): Geostatistical Analysis and Quantification Uncertainty for 3D modelling by Simulation. MSc thesis, ITC, Enschede, The Netherlands.

Bak, R. & Mill, A. (1989): Three dimensional representation in a Geo-scientific Resource Management System for the mineral industry. In: Raper, J. (ed.): GIS – Three dimensional applications in geographic information systems. Taylor&Francis, London.

Breunig, M. & Zlatanova, S. (2006): 3D Geo-DBMS. In: Zlatanova, S. and Prosperi, D. (eds.): Large-scale 3D data integration – Challenges and Opportunities. Taylor&Francis, London, pp. 87-115.

Clark, I. (1979): Practical Geostatistics. Applied Science Publishers, London.

Corbea Diaz, P.P. (1996): Modelling and Visualization of 3D Geo-Spatial Data. MSc thesis, ITC, Enschede, The Netherlands.

Date, C.J. (1995): An introduction to Database Systems. Addison-Wesley Publishing Company, New York.

Dilo, A. (2006): Representation of and reasoning with vagueness in spatial information – A system for handling vague objects. PhD thesis, ITC, The Netherlands.

Egenhofer, M.J. & Herring, J.R. (1990): A mathematical framework for the definition of topological relationships. *Proceedings of the 4th International Symposium on SDH*. Zurich, Switzerland, pp. 803-813.

Foley, J.D. et al. (1990): Computer Graphics: Principles and Practice. 2nd edition. Addison-Wesley Publishing Company.

Foody, G.M. & Atkinson, P.M. (2002): Uncertainty in Remote Sensing and GIS. John Wiley & Sons Ltd.

Fried, C.C. & Leonard, J.E. (1990): 3D in depth: Petroleum models come in many flavours. *Geobyte*, 5, pp. 27-30.

Harbaugh, J.W. & Merriam, D.F. (1968): Computer applications in stratigraphic analysis. Wiley&Sons, New York.

Houlding, S.W. (1994): Uncertainty, Sampling Control and Risk Assessment. In: Houlding, S.W. (ed.): 3D Geoscience Modeling – Computer Techniques for Geological Characterization. Springer-Verlag, Berlin, pp. 185-200.

ISO 19107:2003 (2003): Geographic Information – Spatial schema. International Organization for Standardization (ISO), 2003.

ISO 19108:2002 (2002): Geographic Information – Temporal schema. International Organization for Standardization (ISO), 2002.

ISO 19113:2002 (2002): Geographic Information – Quality Principles. International Organization for Standardization (ISO), 2002.

ISO 19114:2003 (2003): Geographic Information – Quality Evaluation Principles. International Organization for Standardization (ISO), 2003.

ISO 19115:2003 (2003): Geographic Information – Metadata. International Organization for Standardization (ISO), 2003.

Jones, C.B. (1989): Data structures for three-dimensional spatial information systems in geology. *International Journal of Geographical Information Systems*, 3, pp. 15-31.

Lattuada, R. (2006): Three-dimensional representations and data structures in GIS and AEC. In: Zlatanova, S. and Prosperi, D. (eds.): Large-scale 3D data integration – Challenges and Opportunities. Taylor&Francis, London,

pp. 57-86.
Loudon, T.V. (1986): Digital spatial models and geological maps. In: Blakemore, M. (ed.): Proceedings of Auto-Carto Vol. 2, September 14-19. Royal Institution of Chartered Surveyors, London. pp. 60-66.
Lucieer, A. (2004): Uncertainties in Segmentation and their Visualisation. PhD thesis, ITC, The Netherlands.
Meier, A. (1986): Applying relational database techniques to solid modelling. *Computer Aided Design*, 18, pp. 319-326.
Molenaar, M. (1989): An Introduction to the theory of spatial objects modelling. Taylor&Francis, London.
Molenaar, M. (1990): A formal data structure for 3D vector maps. In: Proceedings of EGIS'90 Vol. 2., Amsterdam, The Netherlands, pp. 770-781.
Muller, J.P. (1988): Digital image processing in remote sensing. Taylor&Francis, London.
NEN 3610 (2005): Basic scheme for geo-information - Terms, definitions, relations and general rules for the interchange of information of spatial objects related to the earth's surface. Nederlands Normalisatie-instituut (NEN), 2005.
OGC (1999): OpenGIS Consortium. URL: http://www.opengis.org
OGC (2004): OpenGIS Consortium. URL: http://www.opengis.org
OGC (2006): OpenGIS Consortium. URL: http://www.opengis.org
Oosterom, P.J.M. van; Vertegaal, W.; Hekken, M. van & Vijlbrief, T.: Integrated 3D Modelling within a GIS. International GIS workshop AGDM'94 (Advanced Geographic Data Modelling), Delft, The Netherlands, September 12-14, pp. 80-95, 1994.
Oosterom, P.J.M. van; Stoter, J., Quak, W. & Zlatanova, S. (2002): The Balance between Geometry and Topology. In: Richardson, D. & Oosterom, P.J.M. van (eds.): Advances in Spatial Data Handling. 10th International Symposium on Spatial Data Handling, 2002, pp. 209-224.
Oosterom, P.J.M. van (2003): Vision for the next decade of GIS technology – A research agenda for the TU Delft, the Netherlands. GISt Report No. 20, Delft, The Netherlands, 56pp.
Oosterom, P.J.M. van; Stoter, J. & Jansen, E. (2006): Bridging the worlds of CAD and GIS. In: Zlatanova, S. and Prosperi, D. (eds.): Large-scale 3D data integration – Challenges and Opportunities. Taylor&Francis, London, pp. 9-36.
Orlic, B. (1997): Predicting subsurface conditions for geotechnical modelling. PhD thesis, ITC, The Netherlands.
Pilouk, M. (1996): Integrated modelling for 3D GIS. PhD thesis, ITC, The Netherlands.
Raper, J. (1989): GIS – Three dimensional applications in geographic information systems. Taylor&Francis, London.
Raper, J. & Maguire, D.J. (1992): Design models and functionality in GIS. *Computers & Geosciences*, 18, pp. 387-394.
Requicha, A. (1980): Representations for rigid solids: theory, methods, and systems. *ACM Computing Surveys*, 12, pp. 437-464.
Sides, E.J. (1992): Reconciliation studies and reserve estimation. In: Annels, A.E. (ed.): Case histories and methods in mineral resource evaluation. Geological Society Special Publication, 63, pp. 197-218.
Smets, P. (1991): Varieties of ignorance and the need for well-founded theories. *Information Sciences*, pp. 135-144.
Smets, P. (1996): Imperfect information: Imprecision, and uncertainty. *Uncertainty Management in Information Systems*, pp. 225-254.

Stoter, J. (2004): 3D Cadastre. PhD thesis, Delft, The Netherlands.
Watts, A. (1993): 3D Computer Graphics. Addison-Wesley, Wokingham, England.
Worboys, M.F. (1998): Data Quality in Geographic Information. Hermes-Verlag.
Zadeh, L.A. (1965): Fuzzy sets. *Information and Control*, 8, pp. 335-353.
Zhang, J. & Goodchild, M. (2002): Uncertainty in Geographical Information. Taylor & Francis, London.
Zlatanova, S. (2000): 3D GIS for urban development. PhD thesis, ITC, The Netherlands.
Zlatanova, S.; Rahman, A.A. & Pilouk, M. (2002): Trends in 3D GIS Development. *Journal of Geospatial Engineering*, 4, pp. 1-10.
Zlatanova, S.; Rahman, A.A. & Shi, W. (2004): Topological models and frameworks for 3D spatial objects. *Journal of Computers & Geosciences*, 30, pp. 419-428.
Zlatanova, S. & Stoter, J. (2006): The role of DBMS in the new generation GIS architecture. In: Rana, S. & Sharma, J. (eds.): Frontiers of Geographic Information Technology. Springer-Verlag.

Glossary

Accuracy	Accuracy is the degree of conformity of a measured or calculated quantity to its actual (true) value
Error	An error is a bound on the precision and accuracy of the result of a measurement; classification into two types: statistical error and systematic error; statistical error is caused by random (and therefore inherently unpredictable) fluctuations in the measurement apparatus, systematic error is caused by an unknown but nonrandom fluctuation
Geo-information	Various types of information collected and (re-)used throughout the development of an infrastructural project; i.e. GIS, CAD, multi-dimensional representations; interpretations and representations of surface and subsurface conditions, construction design representations
Harmonization	The act or state of agreeing or conforming; i.e. the process of developing a common way for the management of the various types of geo-information in infrastructural projects, agreeing on common data types and structures
Interpretation Uncertainty	Uncertainties introduced into the representation by the geology experts themselves; a dominant source of uncertainty, especially in geological (subsurface) representations; mainly caused by limitations in data quality and quantity and largely influenced by the knowledge and capability of the interpreter
Precision	Precision characterizes the degree of mutual agreement among a series of individual measurements, values, or results
Quality	The Quality of a product or service refers to the perception of the degree to which the product or service meets the customer's expectations; quality has no specific meaning unless related to a specific function and/or

	object; quality is a perceptual, conditional and somewhat subjective attribute
Uncertainty	Measure of the difference between estimation and reality

Reports published before in this series:

1. GISt Report No. 1, Oosterom, P.J. van, Research issues in integrated querying of geometric and thematic cadastral information (1), Delft University of Technology, Rapport aan Concernstaf Kadaster, Delft 2000, 29 p.p.
2. GISt Report No. 2, Stoter, J.E., Considerations for a 3D Cadastre, Delft University of Technology, Rapport aan Concernstaf Kadaster, Delft 2000, 30.p.
3. GISt Report No. 3, Fendel, E.M. en A.B. Smits (eds.), Java GIS Seminar, Opening GDMC, Delft 15 November 2000, Delft University of Technology, GISt. No. 3, 25 p.p.
4. GISt Report No. 4, Oosterom, P.J.M. van, Research issues in integrated querying of geometric and thematic cadastral information (2), Delft University of Technology, Rapport aan Concernstaf Kadaster, Delft 2000, 29 p.p.
5. GISt Report No. 5, Oosterom, P.J.M. van, C.W. Quak, J.E. Stoter, T.P.M. Tijssen en M.E. de Vries, Objectgerichtheid TOP10vector: Achtergrond en commentaar op de gebruikersspecificaties en het conceptuele gegevensmodel, Rapport aan Topografische Dienst Nederland, E.M. Fendel (eds.), Delft University of Technology, Delft 2000, 18 p.p.
6. GISt Report No. 6, Quak, C.W., An implementation of a classification algorithm for houses, Rapport aan Concernstaf Kadaster, Delft 2001, 13.p.
7. GISt Report No. 7, Tijssen, T.P.M., C.W. Quak and P.J.M. van Oosterom, Spatial DBMS testing with data from the Cadastre and TNO NITG, Delft 2001, 119 p.
8. GISt Report No. 8, Vries, M.E. de en E. Verbree, Internet GIS met ArcIMS, Delft 2001, 38 p.
9. GISt Report No. 9, Vries, M.E. de, T.P.M. Tijssen, J.E. Stoter, C.W. Quak and P.J.M. van Oosterom, The GML prototype of the new TOP10vector object model, Report for the Topographic Service, Delft 2001, 132 p.
10. GISt Report No. 10, Stoter, J.E., Nauwkeurig bepalen van grondverzet op basis van CAD ontgravingsprofielen en GIS, een haalbaarheidsstudie, Rapport aan de Bouwdienst van Rijkswaterstaat, Delft 2001, 23 p.
11. GISt Report No. 11, Geo DBMS, De basis van GIS-toepassingen, KvAG/AGGN Themamiddag, 14 november 2001, J. Flim (eds.), Delft 2001, 37 p.
12. GISt Report No. 12, Vries, M.E. de, T.P.M. Tijssen, J.E. Stoter, C.W. Quak and P.J.M. van Oosterom, The second GML prototype of the new TOP10vector object model, Report for the Topographic Service, Delft 2002, Part 1, Main text, 63 p. and Part 2, Appendices B and C, 85 p.
13. GISt Report No. 13, Vries, M.E. de, T.P.M. Tijssen en P.J.M. van Oosterom, Comparing the storage of Shell data in Oracle spatial and in Oracle/ArcSDE compressed binary format, Delft 2002, .72 p. (Confidential)
14. GISt Report No. 14, Stoter, J.E., 3D Cadastre, Progress Report, Report to Concernstaf Kadaster, Delft 2002, 16 p.
15. GISt Report No. 15, Zlatanova, S., Research Project on the Usability of Oracle Spatial within the RWS Organisation, Detailed Project Plan (MD-NR. 3215), Report to Meetkundige Dienst – Rijkswaterstaat, Delft 2002, 13 p.
16. GISt Report No. 16, Verbree, E., Driedimensionale Topografische Terreinmodellering op basis van Tetraëder Netwerken: Top10-3D, Report aan Topografische Dienst Nederland, Delft 2002, 15 p.
17. GISt Report No. 17, Zlatanova, S. Augmented Reality Technology, Report to SURFnet bv, Delft 2002, 72 p.
18. GISt Report No. 18, Vries, M.E. de, Ontsluiting van Geo-informatie via netwerken, Plan van aanpak, Delft 2002, 17p.
19. GISt Report No. 19, Tijssen, T.P.M., Testing Informix DBMS with spatial data from the cadastre, Delft 2002, 62 p.
20. GISt Report No. 20, Oosterom, P.J.M. van, Vision for the next decade of GIS technology, A research agenda for the TU Delft the Netherlands, Delft 2003, 55 p.
21. GISt Report No. 21, Zlatanova, S., T.P.M. Tijssen, P.J.M. van Oosterom and C.W. Quak, Research on usability of Oracle Spatial within the RWS organisation, (AGI-GAG-2003-21), Report to Meetkundige Dienst – Rijkswaterstaat, Delft 2003, 74 p.
22. GISt Report No. 22, Verbree, E., Kartografische hoogtevoorstelling TOP10vector, Report aan Topografische Dienst Nederland, Delft 2003, 28 p.
23. GISt Report No. 23, Tijssen, T.P.M., M.E. de Vries and P.J.M. van Oosterom, Comparing the storage of Shell data in Oracle SDO_Geometry version 9i and version 10g Beta 2 (in the context of ArcGIS 8.3), Delft 2003, 20 p. (Confidential)
24. GISt Report No. 24, Stoter, J.E., 3D aspects of property transactions: Comparison of registration of 3D properties in the Netherlands and Denmark, Report on the short-term scientific mission in the CIST – G9 framework at the Department of Development and Planning, Center of 3D geo-information, Aalborg, Denmark, Delft 2003, 22 p.
25. GISt Report No. 25, Verbree, E., Comparison Gridding with ArcGIS 8.2 versus CPS/3, Report to Shell International Exploration and Production B.V., Delft 2004, 14 p. (confidential).
26. GISt Report No. 26, Penninga, F., Oracle 10g Topology, Testing Oracle 10g Topology with cadastral data, Delft 2004, 48 p.
27. GISt Report No. 27, Penninga, F., 3D Topography, Realization of a three dimensional topographic terrain representation in a feature-based integrated TIN/TEN model, Delft 2004, 27 p.
28. GISt Report No. 28, Penninga, F., Kartografische hoogtevoorstelling binnen TOP10NL, Inventarisatie mogelijkheden op basis van TOP10NL uitgebreid met een Digitaal Hoogtemodel, Delft 2004, 29 p.

29. GISt Report No. 29, Verbree, E. en S.Zlatanova, 3D-Modeling with respect to boundary representations within geo-DBMS, Delft 2004, 30 p.
30. GISt Report No. 30, Penninga, F., Introductie van de 3e dimensie in de TOP10NL; Voorstel voor een onderzoekstraject naar het stapsgewijs introduceren van 3D data in de TOP10NL, Delft 2005, 25 p.
31. GISt Report No. 31, P. van Asperen, M. Grothe, S. Zlatanova, M. de Vries, T. Tijssen, P. van Oosterom and A. Kabamba, Specificatie datamodel Beheerkaart Nat, RWS-AGI report/GIST Report, Delft, 2005, 130 p.
32. GISt Report No. 32, E.M. Fendel, Looking back at Gi4DM, Delft 2005, 22 p.
33. GISt Report No. 33, P. van Oosterom, T. Tijssen and F. Penninga, Topology Storage and the Use in the context of consistent data management, Delft 2005, 35 p.
34. GISt Report No. 34, E. Verbree en F. Penninga, RGI 3D Topo - DP 1-1, Inventarisatie huidige toegankelijkheid, gebruik en mogelijke toepassingen 3D topografische informatie en systemen, 3D Topo Report No. RGI-011-01/GISt Report No. 34, Delft 2005, 29 p.
35. GISt Report No. 35, E. Verbree, F. Penninga en S. Zlatanova, Datamodellering en datastructurering voor 3D topografie, 3D Topo Report No. RGI-011-02/GISt Report No. 35, Delft 2005, 44 p.
36. GISt Report No. 36, W. Looijen, M. Uitentuis en P. Bange, RGI-026: LBS-24-7, Tussenrapportage DP-1: Gebruikerswensen LBS onder redactie van E. Verbree en E. Fendel, RGI LBS-026-01/GISt Rapport No. 36, Delft 2005, 21 p.
37. GISt Report No. 37, C. van Strien, W. Looijen, P. Bange, A. Wilcsinszky, J. Steenbruggen en E. Verbree, RGI-026: LBS-24-7, Tussenrapportage DP-2: Inventarisatie geo-informatie en -services onder redactie van E. Verbree en E. Fendel, RGI LBS-026-02/GISt Rapport No. 37, Delft 2005, 21 p.
38. GISt Report No. 38, E. Verbree, S. Zlatanova en E. Wisse, RGI-026: LBS-24-7, Tussenrapportage DP-3: Specifieke wensen en eisen op het gebied van plaatsbepaling, privacy en beeldvorming, onder redactie van E. Verbree en E. Fendel, RGI LBS-026-03/GISt Rapport No. 38, Delft 2005, 15 p.
39. GISt Report No. 39, E. Verbree, E. Fendel, M. Uitentuis, P. Bange, W. Looijen, C. van Strien, E. Wisse en A. Wilcsinszky en E. Verbree, RGI-026: LBS-24-7, Eindrapportage DP-4: Workshop 28-07-2005 Geo-informatie voor politie, brandweer en hulpverlening ter plaatse, RGI LBS-026-04/GISt Rapport No. 39, Delft 2005, 18 p.
40. GISt Report No. 40, P.J.M. van Oosterom, F. Penninga and M.E. de Vries, Trendrapport GIS, GISt Report No. 40 / RWS Report AGI-2005-GAB-01, Delft, 2005, 48 p.
41. GISt Report No. 41, R. Thompson, Proof of Assertions in the Investigation of the Regular Polytope, GISt Report No. 41 / NRM-ISS090, Delft, 2005, 44 p.
42. GISt Report No. 42, F. Penninga and P. van Oosterom, Kabel- en leidingnetwerken in de kadastrale registratie (in Dutch) GISt Report No. 42, Delft, 2006, 38 p.
43. GISt Report No. 43, F. Penninga and P.J.M. van Oosterom, Editing Features in a TEN-based DBMS approach for 3D Topographic Data Modelling, Technical Report, Delft, 2006, 21 p.
44. GISt Report No. 44, M.E. de Vries, Open source clients voor UMN MapServer: PHP/Mapscript, JavaScript, Flash of Google (in Dutch), Delft, 2007, 13 p.


ZUIDPOLDER
VAN DELFGAUW
MOLENSLOOT
PARALLELWEG
AKKERDIJKSCHE
POLDER