

# IDENTIFYING THE PROBLEM OF UNCERTAINTY DETERMINATION AND COMMUNICATION IN INFRASTRUCTURAL DEVELOPMENT

Wiebke Tegtmeier <sup>a,b</sup>, Robert Hack <sup>a</sup>, Sisi Zlatanova <sup>b</sup>, Peter van Oosterom <sup>b</sup>

<sup>a</sup> International Institute for Geo-Information Science and Earth Observation (ITC), Hengelosestraat 99, 7500 AA Enschede, The Netherlands – (tegtmeier, hack)@itc.nl

<sup>b</sup> University of Technology Delft (TU Delft), Onderzoeksinstituut OTB, Jaffalaan 9, 2628 BX Delft, The Netherlands – (w.tegtmeier, s.zlatanova, p.j.m.vanoosterom)@tudelft.nl

**KEY WORDS:** Quality, Uncertainty, Interpretation Uncertainty, Geo-Information, Real World Representations, Geoportals

## ABSTRACT:

In civil engineering infrastructural projects many different parties are involved into different phases during the lifetime of a construction; that are consultants, contractors, architects, suppliers of subsurface and surface data, governmental bodies, politicians, the client himself, etc. The communication between these different parties is difficult. This is essentially 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 process. This leads to the introduction of “interpretational uncertainties” into the representation; that are, uncertainties introduced by the expert while using own knowledge and experience for the data interpretation. To determine possible ways of data quality and uncertainty handling in subsurface representations, interviews with professionals were carried out and companies (engineering/ consulting/ insurance companies, etc.) visited. Further, a number of Dutch geoportals were checked regarding their provision of quality and uncertainty information. This paper exposes the influence of the expert knowledge on the quality and uncertainty of a subsurface real world representation. It also summarizes the results of visits and interviews as well as the content of the various geoportals visited. Common approaches for the implementation of quality and uncertainty expressions in subsurface real world representations are presented. Finally, an outlook is given on the planned development of methodologies to describe the “interpretational uncertainty” in quality and uncertainty expressions and to include these, together with additional information, in the metadata of the subsurface real world representation.

## 1. INTRODUCTION

Civil engineering infrastructural development is becoming more and more important in order to provide the required mobility to people all around the world. To be able to cover the needs of the world’s citizens, sufficient infrastructural capacities (e.g. highways, railways, waterways, airports, etc.) must be made available. In the same way, infrastructure security and transportation safety must be ensured. Clearly, this requires the cooperation and clear communication of a number of specialists, each of them dealing with different types of problems. These different specialists may be, for example, consultants, contractors, architects, suppliers of subsurface and surface data, governmental bodies, politicians, the client himself, etc. New civil engineering infrastructure must be planned, designed, and built and existing structures monitored, maintained and eventually be abandoned. These diverse processes are commonly spread over the whole lifecycle of civil infrastructures, throughout which large quantities of geo-information (e.g. GIS-, CAD-, and other data sets) are collected, generated and (re-) used.

As known today, problems in infrastructural development are often connected to the exchange and (re-) use of geo-information and are largely caused by a lack of quality and uncertainty information. Since large parts of the collected data and retrieved real world representations are not equipped with individual information about data origin, collection method, quality and possible uncertainties, the communication between the different professionals is difficult and, thus, the (re-) use of the information aggravated. This problem needs to be solved in order to improve cooperation between the numerous companies

involved in infrastructural development and to make the process of infrastructural development more cost-effective. Consequently, the following question arises: ‘How can quality and uncertainty in real world representations be determined, defined and finally included in the metadata tables of these representations?’

## 2. THE QUALITY ASPECT OF GEO-INFORMATION

Regarding the fact that numerous parties involved in infrastructural development make use of available geo-information in order to take decisions, the quality of this information is of major importance. Quality is therefore an important aspect of geo-information and, to allow an effective use of collected data and information, it is necessary to know its quality (Hack, 1997; Dilo, 2006). To be able to determine and also to define the quality of various types of geo-information, it is important to understand what the word “quality” actually means.

Many definitions of quality exist in the literature. Despite this large number of quality definitions available in the literature, especially scientists and engineers are, still, frequently using the definition of quality as defined in the ISO standards (e.g. ISO 9001:2000). This definition of quality 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. This is not to be mistaken for ‘degree of excellence’ or ‘fitness for use’ that meets only part of the definition.”

As described by a number of scientists (e.g. Smets, 1996; Dilo, 2006), data quality can be influenced by various factors that may lead to different types of imperfect data (Dilo, 2006). According to Smets (1996), imperfect data can have diverse aspects, such as imprecision, inconsistency, and uncertainty. With it, imprecision and inconsistency are properties of the data, whereas uncertainty is introduced into the data by attaching weights to the worlds in order to express our opinion about which might be the real world situation.

Considering the complexity of the problem of quality in geo-information, especially with respect to the various factors influencing quality and causing different types of imperfections, it has been decided that this research will foremost be focused on the quality aspect of uncertainty. With it, the determination and definition of uncertainties in geo-information and real world representations of the subsurface are of greatest concern, since for the interpretation of subsurface conditions usually limited information is available.

## 2.1 Uncertainty as part of the quality aspect

It is important to realize that uncertainty plays an important role in the development of infrastructural projects. The main, but often neglected, problem is that uncertainty creeps into infrastructural development through every pore. Whether geo-information is collected, processed, or interpreted; thus, whether experts perform site investigations, select interpretation procedures and modelling parameters or give indications about real world conditions, they are usually walking on very slippery terrain. Thereby, it is difficult for non-experts - and sometimes even for other experts - to appreciate how complex these tasks are and how easy it is for different people to come to different conclusions. Since numerous companies are involved in different phases of the process of civil engineering infrastructural development, possible uncertainties in the geo-information or any of the real world representations must be properly defined and communicated to ensure an unhindered cooperation between the different parties during infrastructural development. Despite an apparent need to reduce the uncertainty from an end-users and decision-makers perspective, however, it is, still, not possible to eliminate this factor of uncertainty (Foody & Atkinson 2002).

Uncertainty generally exists when details of situations are ambiguous, complex, unpredictable, or probabilistic (Brashers, 2001). Further, uncertainty develops when information is unavailable or inconsistent, and when people feel insecure in their own state of knowledge or the state of knowledge in general (e.g. Babrow, Kasch & Ford, 1998; Babrow, Hines & Kasch, 2000). Uncertainty itself is a rather indistinct theme, which is often described as '...a measure of the difference between estimation and reality...'. That can be, for example, the difference of the thickness of a soil layer as determined with CPTs compared to the situation in reality; expressed in percentage. This, somewhat broad, 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'.

In the same way as quality, which is influenced by various factors causing different types of data imperfection (e.g. uncertainty), also the quality aspect of uncertainty itself proves to be affected by a number of different factors. In line with the conceptualisation of sources of uncertainty for physicians, given

by Beresford (1991), uncertainty in geo-information can be categorised into three types; that are namely technical, personal, and conceptual uncertainties. With it, different types of uncertainty can be determined, that are, for example, uncertainty resulting from site investigations, surveys, or measurements (i.e. technical uncertainty), uncertainty resulting from geological and geotechnical interpretations (i.e. personal uncertainty) or uncertainty with regard to spatial prediction (i.e. conceptual uncertainty). According to Beresford (1991), technical sources of uncertainty are often defined as 'first order uncertainties', whereas personal and conceptual sources of uncertainty are also defined as 'second order or meta-uncertainties' and refer to the uncertainty about how uncertain one is. What makes Beresford's analysis particularly useful is that it shows that the management of uncertainty is more complicated than the simple provision of more information. Further information would merely address the technical or first order sources of uncertainty, and leave untouched meta-uncertainty (Hall, 2002).

Accordingly, the many techniques (e.g. geostatistical simulations, kriging, or probability-based methods) available today for the determination of (first order) uncertainty resulting from the process of spatial prediction (Orlic 1997, Zhang & Goodchild 2002) prove not to be enough. Alas, the uncertainty associated with spatial prediction is, by far, not the only factor of uncertainty to be taken into consideration in the process of infrastructural development. At least two other main sources of uncertainty exist that one should be aware of throughout the different phases of infrastructural development. That are, on the one hand, the potential for investigation errors (e.g. location errors or measurement errors caused by wrongly calibrated machines) and, on the other hand, 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). Although these uncertainties are recognized and more or less described, it is still difficult to quantify and define some of these sources of uncertainty. Regardless of the apparent need to quantify the uncertainty associated with investigation and interpretation errors, there is at present little what can be done about these specific types of uncertainty in geo-information.

Clearly, more research is needed in this direction, especially regarding the influence of the so-called 'interpretation uncertainties' on the process of infrastructural development. Interpretation uncertainties can be described as uncertainties introduced into the representation by the geology or geotechnical experts, e.g. by applying own knowledge and experience for the data interpretation. These interpretation uncertainties are, especially in geological (subsurface) representations, a dominant source of uncertainty and need to be further defined and properly communicated.

In order to further clarify the problem of interpretation uncertainties, a small example will be given that describes the case of marine and fluvial settings (see also Figure 1).

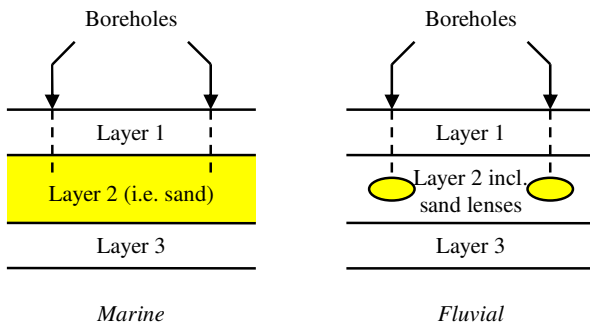


Figure 1. The case of marine or fluvial settings – The interpretation outcome as a result of available data and a priori knowledge of the specialist

In the western part of the Netherlands most sedimentary layers have either a marine or fluvial origin. Assume that a foundation has to be made on a sand body in the subsurface. Some boreholes have been made and all show a sand layer to exist roughly at the required depth. Now the interpretation starts. If the sand layer is of marine origin, it can be assumed with reasonably safety that the layer is continuous. However, if the sand layer is of fluvial origin, it is very likely to be a lens with a limited lateral extension, and may or may not be continuous between the two boreholes (see also Hack et al., 2006). That is the point where the knowledge and experience of the engineer/geologists comes in and possible interpretation uncertainties might be introduced into the subsurface representation. The specialist, who knows in which formation the sand layer is situated (i.e. of marine or fluvial origin), will most likely make a correct interpretation, while his colleague, who does not know the specific geological settings or who makes the wrong assumptions, may produce a completely wrong interpretation; with all consequences for the foundation and the building resting on it.

What should become clear by means of this example is that interpretation uncertainties are far more than just erroneous measurements. They might be caused or increased by measurement errors; however, they can also exist even though the quality of the measurements is good, since they are caused during the interpretation process of the data and largely dependent on the experience and a priori knowledge of the interpreter. According to Hack et al. (2006), the quality of this a priori knowledge that is essential in the interpretation can, at present, not be quantified in general. If the engineer/geologist is good, the interpretation of the data will be good; if the engineer/geologist is not so good, it will result in a poor data interpretation. The establishment of, for example, rock or soil layer boundaries or boundaries of other geological features depends, thus, largely on engineering judgement.

Regarding this and the fact that for the interpretation of conditions below the earth's surface limited information is available, parts of the research are focused on the determination and communication of 'interpretation uncertainties' in subsurface real world representations. For the proper definition and communication of uncertainties associated with the interpretation process during various phases of infrastructural development, certain steps are followed throughout this research. First, a number of geoportals frequently used for data

and knowledge exchange in the Netherlands are investigated and, foremost, checked regarding their content of quality and uncertainty indications. Second, interviews (i.e. in form of a questionnaire) are taken out with (engineering) companies, governmental bodies, etc. in order to receive information about their way of quality and uncertainty handling in infrastructural projects. In addition, the topics of geo-information collection, use and exchange are covered in the questionnaire. Finally, a first step will be made towards the determination of the so-called 'interpretation uncertainties' in subsurface real world representations. These various steps and their (preliminary) outcome will now be focused at in the coming sections.

### 3. GEOPORTALS

Geoportals can generally be described as 'World Wide Web gateways that organize content and services such as directories, search tools, community information, support resources, data and applications' (Maguire & Longley, 2005). The word 'portal' stems from the Latin word porta and indicates an entrance point (Annoni et al., 2004). In line with Annoni's definition, Tait (2005), more specifically, describes a geoportal as 'a web site considered to be an entry point to geographic content on the web or, more simply, a web site where geographic content can be discovered'.

In infrastructural development, geoportals can be useful tools when trying to achieve easier communication between the various parties. Via geoportals, the geo-information itself, but also documentation (in hard copy or more frequently in digital form) that describes the provenance, ownership, quality, age, fitness for purpose and other useful properties can be easily made accessible. This 'associated' data documentation, or data about data, which is often referred to as metadata, plays an important role regarding the improvement of geo-information exchange and (re-) use. According to Maguire & Longley (2005), geoportals provide capabilities to query such metadata records for relevant data and services using space, time, and thematic attributes (see also Figure 2; illustration of a query based on space and also thematic attributes), and then link directly to the on-line content services themselves. In addition, they can control the commercial usage of services by facilitating the sale/purchase of data and services.

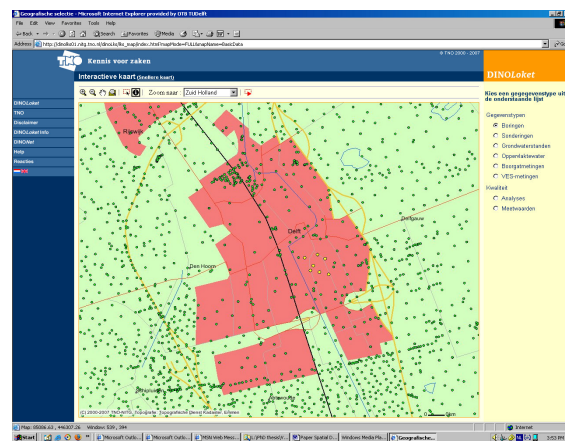


Figure 2. Search page from the DinoLoket portal showing the 'zoom in'-window for location based search (www.dinoloket.nl)

Throughout the last years, the use of geoportals is significantly increased. A large number of geoportals have been developed and taken into use; that is in the Netherlands just as all around the world. These geoportals have proven to facilitate notably the access and exchange of the various types of geo-information. Via geoportals, geo-information can easily be added, extracted, but also deleted and updated. Contractors can share experience from site investigations and future construction companies can obtain important knowledge in order to, again, make their own working processes more cost-effective, and improve the safety of co-workers but also of the final construction itself.

In this research, a number of Dutch geoportals are investigated and compared concerning their content, access, search mechanism, functionality, pricing, etc. Special interest is given to the implementation of quality and/or uncertainty indications regarding geo-information and real world representations in the geoportals.

### 3.1 Geoportals in the Netherlands

Nine Dutch geoportals are studied in total. These geoportals contain various types of (geo-) information that are considered relevant for infrastructural development; such as geological, geotechnical data, environmental information, digital maps, satellite images, roadways, waterways, railways, urban information, etc.

Depending on the geoportal used, information can be added via mail, e-mail, online (e.g. directly uploaded or inserted via the geoportal), fax, etc. In the same way, information can be derived via mail, e-mail, online, fax etc. In addition, the search mechanisms vary depending on the geoportal visited. Via the various geoportals, the user can discover which service is available regarding a particular topic (e.g. CPT data, housing prices, and cadastre data), geographic area, time period, etc. Further, a number of different file formats are available; frequently used are, however, ASCII, .pdf, .shp, etc.

The Dutch geoportals visited throughout this research together with their content and important characteristics are the following:

- Het rode loket (run by DataLand and the Dutch Central Bureau of Statistics):

Content: building information

Information: location (i.e. address, x-/y- coordinates), construction-specific (i.e. volume, area, no. of floors, year of construction), use (i.e. current use, rented out or owned), additional (i.e. no. of residents, mean value)

Selection: on basis of area, province, city, ZIP code, type of information

Pricing: 0,10-0,30 € depending on requested information (i.e. type and number of data)

Data format: ASCII; information delivered following DataLand specifications (i.e. StUF-DL)

- Het groene loket (run by a large variety of companies specialized in the different fields covered by this geoportal):

Content: various topics (e.g. geology, agriculture, biodiversity, meteorology, environment, cadastre, society, transport, etc.)

Information: content type, update frequency, x-/y- coordinates, spatial data type, ID, language, company contact, metadata standard, distributor, (spatial) data quality, etc.

Selection: on basis of data categories, data types and formats, location, area, place, address, latitude/ longitude, time period, date posted, data provider

Pricing: free of charge

Data format: ArcGIS maps, MS Access tables, .jpg, .pdf etc.

- Het bruine loket (run by a large variety of companies specialized in the different fields covered by this geoportal):

Content: subsurface geological/ geotechnical information

Information: see 'Het groene loket'

Selection: see 'Het groene loket'

Pricing: see 'Het groene loket'

Data format: see 'Het groene loket'

- Het witte loket (initiative of GeoNetwork opensource):

Content: various topics (e.g. biology, biodiversity, forestry, ecosystems, water, vegetation, oceanography, etc.)

Information: ID, citation information (i.e. originator, title, publication date & place, publisher), abstract, time period, status (i.e. progress, maintenance, updates), constraints, metadata reference information (i.e. date, contact, standard)

Selection: on basis of title, abstract, free text, keywords, location, data categories, map type, server

Pricing: depending on data type and deliverer (i.e. type and number of data)

Data format: various datasets, satellite imagery, technical reports, maps

- Geodatabank (run by GeoDelft and Fugro):

Content: CPT and borehole data

Information: ID, type of investigation, detailed method description, begin/end depth, reference height, x-/y- coordinates, accuracy, street, place, ZIP code, year, measurement results

Selection: on basis of location (i.e. using 'zoom in'-function or search function using place, ZIP code, x-/y- coordinates) or search filter (i.e. type of method, street name, begin/end year, length, height reference)

Pricing: 50 € basic costs plus 15-90 € depending on the selected product (i.e. type and number of data)

Data format: .tiff

- GeoBrain (run by GeoDelft):

Content: boring and foundation techniques (i.e. construction sheet pile walls, removal sheet pile walls, construction prefab piles, construction vibro piles)

Information: location information, site information, geotechnical information, construction specific information, installation information, design information, neighbourhood information, investigation systems, positioning systems, experiences

Selection: on basis of foundation or boring technique and location (i.e. using 'zoom in'-function or name of province) or specific method or specific ground structure

Pricing: free of charge

Data format: data inserted via .pdf forms (i.e. using 'Geotechnical Exchange Format GEF'), results online

- DinoLoket (run by TNO-NITG):

Content: (sub-) surface information (i.e. seismic measurements, boreholes, borehole measurements, CPTs, groundwater levels, water quantity & quality, soil measurements, geological & geotechnical maps and profiles, etc.)

Information: ID, x-/y- coordinates, coordinate system, place, province, map sheet, positioning system, investigation specific information (e.g. begin/end depth boring), date, client, contractor, layer description (i.e. lithological description, company, method, date, quality), measurement results

Selection: on basis of type of information measurement and location (i.e. using 'zoom in'-function or name of province) or ID of specific measurement

Pricing: free of charge for private and research purposes; charges for company use

Data format: ASCII, .xml, .pdf (i.e. using GEF- and NEN-standards)

- Kadaster (run by the Dutch Cadastre):

Content: cadastral information

Information: (parcel, house, apartment) types, rights, possessions, mortgages, restrictions, easement on properties, ownership, surface area, prices, year of purchase, location, cadastral map, neighbourhood map, boundary reconstruction, topographic maps, aerial photographs, etc.

Selection: on basis of required information and location

Pricing: depending on requested information (i.e. type and number of data)

Data format: .pdf, hardcopies, .dgn, .dxf, .shp, .bmp, .tiff, ASCII, etc.

- Rijkswaterstaat (run by the Dutch ministry for traffic and waterways)

Content: traffic and (water)ways information

Information: topographic maps, height maps, digital terrain models (DTMs), aerial photographs, national base geo-information (i.e. surface level information, coordinate system converter), geo-information network 'dry' (i.e. topographic data of road infrastructure, deformation measurement, etc.), geo-information network 'wet' (i.e. topographic data of waterways, dams, surface water bodies, etc.)

Selection: on basis of product type, location (i.e. province), map sheet

Pricing: depending on the product; some free of charge, some against small charges (i.e. type and number of data)

Data Format: .dwg, .dxf, .shp, ASCII

### 3.2 Quality and uncertainty in Dutch geoportals

Quality and uncertainty indications are, at present, still largely missing in most Dutch geoportals. As becomes clear when regarding the content and the information provided by the geoportals described earlier in this paper, only a few geoportals start including descriptions of quality and (possible) uncertainties concerning their geo-information.

In 'het bruine loket' and 'het groene loket', a first step is made towards the description of the spatial data quality of the geo-information provided. A so-called 'quality element' is included in the metadata of the geo-information. With it, information regarding the producer of the geo-information as well as the method and date of production are given. Further, the positional accuracy and the completeness of the data set (i.e. expressed in percentage of entities missing) are offered. Unfortunately, not all types of geo-information that are available via these two geoportals are equipped with the quality information. Furthermore, the quality information provided focuses mainly on the spatial aspect and the accuracy of the geo-information. Uncertainty estimations regarding the data interpretation process or information about the experience of the expert undertaking the interpretation process are not given.

Also the 'DinoLoket' portal, which is the central geoportal for data and information of the Dutch subsurface, is starting with the development of so-called 'quality codes' that can be attached to the geo-information and included in the measurement result tables as well as in the geo-information's metadata. These quality codes mark the quality of, for example, a soil layer description, or groundwater quantity and quality measurements. The determination of the quality codes is based, for instance, on the average number of lithological features described per layer or the average number of quality and quantity characteristics described per water filter. For borehole information, quality codes from A to F are defined, meaning the description of the litho logy is 'very detailed' (i.e. A) or 'no layer information available' (i.e. F). For water quality and quantity measurements, quality codes from A to C and O are defined, meaning the description of water quality and quantity is either 'very good/ reliable' (i.e. A), 'sufficient/ reliable' (i.e. C) or 'unsatisfactory/ unreliable' (i.e. O). In addition, the quality code X can be given when the data available is too little for a proper description. What should be noted is that the quality code does not give an indication regarding the quality or (possible) uncertainties of the content of the data interpretation itself. It only defines the extent of the litho logy or water quality and quantity description. In the same way as 'het groene loket' and 'het bruine loket', the 'DinoLoket' does not contain quality codes for all types of geo-information and measurements available at the geoportal. In addition, no further information concerning the quality of the content of the data interpretation or the knowledge and experience of the interpreter is given.

Apparently, quality and uncertainty information is still missing in most geoportals available in the Netherlands. Only few geoportals are starting to provide quality codes as part of their geo-information or metadata; however, only a small number of geo-information types are equipped with these quality codes. Another drawback about these quality codes is also that they give no further information regarding the quality or uncertainty of the content of the information provided. The uncertainties associated with the interpretation process (i.e. depending on the experience and knowledge of the interpreter) are, thus, not considered in the quality codes so far.

## 4. QUALITY AND UNCERTAINTY HANDLING IN DUTCH (ENGINEERING) COMPANIES AND PROVINCES

In addition to the Dutch geoportals visited in the course of this research, a number of Dutch (engineering) companies and provinces known to be frequently involved in infrastructural development projects were approached with a questionnaire. To large parts, this questionnaire covers the topics of semantics of geo-information as well as the use of attribute information and metadata together with geo-information standards; with special interest in common practices in the various companies and provinces. The second part of the questionnaire contains questions regarding the management of quality and uncertainty in infrastructural development; that is, with focus on common practices on inter-company and inter-project levels. The questions regarding the handling of quality and uncertainty information in infrastructural development were the following:

1. In your company, is there any quality control taking place regarding your geo-information (e.g. regarding data, representations & interpretations)?

2. In case there is a quality control taking place, at which level is it applied and how do you determine and describe this quality?
3. Where and in which form in the representation are you storing this quality information?
4. Are you including uncertainty estimations into your representation?
5. In case there is uncertainty estimation taking place, at which level is it applied (e.g. object or representation level) and how do you determine and describe this uncertainty?
6. Where and in which form in the representation are you storing this uncertainty information?
7. Are you using metadata in order to offer information about possible uncertainties as well as the implementations and meaning of the real world representations?
8. In which way is your quality and uncertainty control supported with ICT-instruments?
9. Are you taking into account the influence of 'interpretation uncertainties'; that are uncertainties introduced into the representations by the (geology) expert himself by applying knowledge and experience during the interpretation?
10. If yes, how are you defining and communicating this information?
11. Are you using metadata in order to offer information about interpretation uncertainties regarding real world representations?
12. Are you applying ICT-instruments in order to determine and communicate information about interpretation uncertainties?

A number of companies and provinces were approached with this questionnaire and, until now, 12 questionnaires were returned. In three of the questionnaires, it was indicated that no quality and/ or uncertainty control is taking place, neither on object nor on representation level. Thus, in these companies, no information regarding the quality or (possible) uncertainties of the geo-information is determined and included in the attribute information or metadata of the geo-information. In seven questionnaires, it was indicated that quality control concerning the geo-information is taking place. The quality control is, in these cases, done at representation level. In case the representation process (e.g. of a CAD design/ drawing or GIS representation) is completed, it will be checked by supervisors or co-workers. This process of quality control is, still, for the most part done manually and, hence, without the support of computers. In these companies, the ICT-support for the quality control of geo-information can be considered minimal to zero. Only in one company, that is a company specialized on visualization, the visualization software has a built-in quality control function. The derived quality information is, in the end, generally stored using pre-defined forms (e.g. verification forms) or standards (e.g. NEN), that is digitally at representation level in a project folder, but also as a hardcopy in the company's archives. Alas, none of these companies is concerned with the determination of interpretation uncertainties or use attribute information or metadata for the storage of quality indications. In two questionnaires it was indicated that next to the quality control also the influence of the so-called interpretation uncertainties and, hence, the influence of the expert himself throughout the data interpretation process, is taken into account. In both companies, the quality control is done at representation level using ArcGIS as software support tool as well as standard checklists. The results of the quality control process are included in the metadata of each

representation. The process concerning the determination of interpretation uncertainties is, unfortunately, not further explained in the questionnaire. However, just as the quality indications, the indications regarding (possible) uncertainties resulting from the interpretation process are stored as metadata together with the corresponding representation. Finally, the geo-information as well as attribute information and metadata is, in these two companies, managed in project-specific folders inside the companies.

The exchange of information; that is geo-information in general and also information regarding quality and (possible) uncertainties of the geo-information used; concerning infrastructural projects is on inter-company as well as on inter-project levels, still, mainly taking place orally, via hardcopies or email. Only in very few cases, metadata is used in order to exchange 'associated' data documentation; that is additional data about data, such as quality and uncertainty indications.

In the same way as in the Dutch geoportals, quality and uncertainty information are, thus, often neglected in many (engineering) companies and provinces involved in infrastructural development in the Netherlands. Few companies apply quality control in their companies and even fewer communicate this information to project partners with the help of metadata attached to the specific geo-information types. Only a small number of companies try to determine and communicate the influence of the interpretation uncertainties on the final real world representations as well as on the safety of the final construction.

## **5. CRITICAL RESEARCH ISSUES AND DEVELOPMENTS**

Regarding the fact that a clear and unhindered communication is of major importance in infrastructural development in order to ensure easy exchange and (re-) use of geo-information and to make the construction process cost-efficient and safe, it is vital to equip the geo-information with sufficient indications regarding its quality and (possible) uncertainty. These two factors have proven to be of substance when trying to communicate between different companies in a project and when trying to exchange different types of geo-information. If people have no further information about the origin and quality of data and representations, they will not simply rely on this information and use it for future working steps (Hack et al., 2006).

With respect to the situation in geology and geotechnology, the use of quality and uncertainty indications is even more important. Since for the representation of the subsurface (geotechnical) situation at a construction site only sparse information is available, the knowledge and experience of the interpreter plays an important role regarding the outcome of the interpretation. Despite the importance of these uncertainties in subsurface representations, this topic is, still, to large parts neglected in today's' civil engineering infrastructural projects. Many up-to-date analyses are available and applied describing all sorts of uncertainties (primarily 'first order uncertainties') in measurable properties. Without an indication on the level of interpretation uncertainties to be expected in the representation, it is, however, difficult to rely on any geotechnical representation and to use it for further planning and decision-making.

By means of the geoportal search as well as the questionnaires it has become clear that the determination and communication of quality and uncertainty in geo-information is still sparse concerning the process of infrastructural development. Indications about the quality of geo-information can already be found in a number of geoportals and companies. However, this information is primarily concerning the spatial data quality, that is the locational accuracy, and less the quality of the content of an interpretation. Also, the quality information is hardly integrated in the metadata of the geo-information. Often, quality indications are stored in pre-defined lists in project folders and, additionally, kept as hardcopies in company-own archives. The communication of data and geo-information quality, if existent at all, is primarily done orally and also via mail from department to department inside the companies or from company to company inside projects. As to the determination of (interpretation) uncertainties, especially regarding subsurface real world representations, this is, still, minimal to non-existent.

Nevertheless, what has also become clear throughout the research concerning quality as well as uncertainty handling in geoportals as well as companies and provinces in the Netherlands is that this problem is known to be existent and first steps are made towards a solution. With it, several companies in the same way as the geoportals 'GeoBrain' and 'DinoLoket' are busy with so-called 'quality improvement projects'. These projects are aiming at an improved and easier exchange of geo-information via geoportals. Hereby, geo-information will be checked for its origin, completeness and correctness and will finally be equipped with sufficient metadata, also including quality and uncertainty indications.

## 6. CONCLUSIONS AND FUTURE RESEARCH

What has been emphasized in this article is that a reasonable determination and communication of quality and (possible) uncertainties in real world representations would significantly ease the cooperation between all parties involved in the infrastructural projects. With it, the determination, description and communication (ideally via metadata) of the so-called 'interpretation uncertainties' in subsurface real world representations is of major interest. A suitable method for the quantification and communication of such uncertainties in geotechnical real world representations is thought to be determined throughout this research.

Part of this research will, thus, be focused on the determination and communication of interpretation uncertainties in subsurface geotechnical representations. As this problem cannot be completely solved in this limited amount of time, a first step will be made towards an acceptable solution. Within this research, a method will be developed, which can be used to achieve a description of the level of interpretation uncertainty to be expected in a certain representation of the subsurface. This level of interpretation uncertainty in geotechnical representations is, at present, intended to be described on a scale of, for example, 1 to 5; with "1" representing a low level of interpretation uncertainty and high reliability of the subsurface representation and "5" vice versa. For the determination of the level of interpretation uncertainty, a weighting system will be applied. The use of this weighting system will lead to the derivation of scaled values indicating expected interpretation uncertainties in a certain geotechnical representation as well as their influence on the construction and its safety. Aspects that will be taken into account in the

weighting system are, among others, the quantity of the collected data, the quality of the collected data, the extent/ size of the construction site, the expected impact of the civil construction on the geology (i.e. type/size/etc. of construction) and the experience of the geotechnical expert responsible for the interpretation (i.e. familiarity with geology around the construction site, number of representations made in this area, etc.). Each of these aspects will then be described by a factor. Additionally, these factors are weighted depending on the expected influence on the final interpretation uncertainty regarding the geotechnical representation.

Finally, the newly determined uncertainty information will, ideally, be included in the metadata of the subsurface (geotechnical) representation and, if possible, be equipped with supplementary information regarding the implications of this interpretation on the construction of the infrastructural project. This should significantly improve the communication between the companies involved in infrastructural development, facilitate the (re-) use of the geo-information and, finally, make the construction more cost-efficient and safe.

## REFERENCES

- Annoni, A., Bernard, L., Fullerton, K., de Groof, H., Kanellopoulos, I., Millot, M., Peedell, S., Rase, D., Smits, P. & Vanderhaegen, M., 2004. Towards a European spatial data infrastructure: the INSPIRE initiative. In: *Proceedings of the 7th international global spatial data infrastructure conference*, Bangalore, India, 11pp.
- Babrow, A.S., Hines, S.C. & Kasch, C.R., 2000. Managing uncertainty in illness explanation: An application of problematic integration theory. In: *Explaining illness: Research, theory, and strategies*. Erlbaum, Hillsdale, NJ, pp. 41-67.
- Babrow, A.S., Kasch, C.R. & Ford, L.A., 1998. The many meanings of uncertainty in illness: Toward a systematic accounting. *Health Communication*, 10, pp. 1-23.
- Brashers, D.E., 2001. Communication and Uncertainty Management. *Journal of Communication*, 51(3), pp. 477-497.
- Beresford, E.B., 1991. Uncertainty and the shaping of medical decisions. *Hastings Cent Report*, 21, pp. 6-11.
- Dilo, A., 2006. Representation of and reasoning with vagueness in spatial information—A system for handling vague objects. PhD thesis, ITC, Enschede, The Netherlands.
- Foody, G.M. & Atkinson, P.M., 2002. *Uncertainty in Remote Sensing and GIS*. John Wiley & Sons Ltd, West Sussex, 307 pp.
- Hack H.R.G.K., 1997. Digital data for engineering geology: disaster or benefit? In: European Science Foundation, "Virtual environments for the Geosciences", *Space-time modelling of bounded natural domains*. Rolduc, the Netherlands.
- Hack, R., Orlic, B., Ozmutlu, S., Zhu, S. & Rengers, N., 2006. Three and more dimensional modelling in geo-engineering. *Bulletin of Engineering Geology and the Environment*, 65, pp.143-153.

Hall, K.H., 2002. Reviewing intuitive decision-making and uncertainty: the implications for medical education. *Medical Education*, 36(3), pp. 216-224.

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

<http://dinolks01.nitg.tno.nl/dinoLks/DINOLoket.jsp>

<http://www.geobrain.nl/>

<http://www.geodatabank.nl/intro.html>

<http://www.geoloketten.nl/jahia/Jahia/cache/offonce/lang/en/pid/288;jsessionid=25A9A45ED449E9B68FAABD00DEAB7A60>

<http://www.kadaster.nl/>

<http://www.rws.nl/rws/agi/home/html/menu2/geo-informatie.htm?home>

ISO 9001:2000 Quality Management Standard.

Lattuada, R., 2006. Three-dimensional representations and data structures in GIS and AEC. In: *Large-scale 3D data integration-Challenges and Opportunities*. Taylor&Francis, London, pp. 57-86.

Maguire, D.J. & Longley, P.A., 2005. The emergence of geoportals and their role in spatial data infrastructures. *Computers, Environment and Urban Systems*, 29(1), pp. 3-14.

Orlic, B., 1997. Predicting subsurface conditions for geotechnical modelling. PhD thesis, ITC, Enschede, The Netherlands.

Smets, P., 1996. Imperfect information: Imprecision, and uncertainty. *Uncertainty Management in Information Systems*, pp. 225-254.

Tait, M.G., (2005). Implementing geoportals: applications of distributed GIS. *Computers, Environment and Urban Systems*, 29(1), pp. 33-47.

Zhang, J. & Goodchild, M., 2002. *Uncertainty in Geographical Information*. Taylor & Francis, London.