GIS TECHNOLOGY

Geo-information and formal semantics for disaster management

PhD Research Proposal

Wei Xu, MSc

GISt Report No. 46

December 2007

TU Delft

OTB Research Institute for Housing, Urban and Mobility Studies

Delft University of Technology

Geo-information and formal semantics for disaster management

PhD Research Proposal

Wei Xu, MSc

GISt Report No. 46 December 2007

Summary

This proposal contains the research plan for the PhD research "Geo-information and Formal Semantics for Disaster Management". It describes relevant backgrounds, goals, phasing, research questions and the planning. The research will be carried out in four-year period during 2006-2010. This research is closely related to a Dutch project GDI4DM (Geo-spatial Data Infrastructure for Disaster Management) and European project Humboldt. This PhD project will investigate the usage of ontology to formally describe the semantics of geo information for disaster management. It is the assumption that with the help of ontology the semantics of geo information can be explicitly expressed in order to enable machine-automation for disaster management. Questions such as "what is the added value of formal semantics for geo-information", "How can formal semantics be used" will be answered after the PhD research.

ISBN: 978-90-77029-22-0
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: wei.xu@tudelft.nl

Table of Contents

Abstract

This proposal contains the research plan for the PhD research “Geo-information and Formal Semantics for Disaster Management”. It describes relevant backgrounds, goals, phasing, research questions and the planning. The research will be carried out in four-year period during 2006-2010. This research is closely related to a Dutch project GDI4DM (Geo-spatial Data Infrastructure for Disaster Management) and European project Humboldt. This PhD project will investigate the usage of ontology to formally describe the semantics of geo information for disaster management. It is the assumption that with the help of ontology the semantics of geo information can be explicitly expressed in order to enable machine-automation for disaster management. Questions such as ”what is the added value of formal semantics for geo-information”, ”How can formal semantics be used” will be answered after thePhD research.

Chapter 1: **Introduction**

1.1 Background

Increasing numbers of natural and man-made disasters, such as earthquakes, tsunamis, floods, air crashes, have posed a challenge to public services and demonstrated the importance of disaster management. The five phases of disaster management, namely mitigation, prevention, preparedness, response and recovery [45], have urged collaboration among various users such as the Fire Brigade, Police, Medical Services, the Municipality (local authorities), Red Cross, urban planners and even amongst different countries.

Research is carried out and large investments are made in the area of disaster management (the list is just an indication but not an exhaustive one). For example, the following projects range from the preparedness to disaster recovery (a more complete list is available on `www.psc.europe.eu`).

- DIPECHO (DIsaster Preparedness European Commission's Humanitarian aid department) is an European Union's program which "supports training, capacity-building, awareness-rasing and early-warning" (website: `http://ec.europa.eu/echo/field/dipecho/index_en.htm`).
- EU-MEDIN (Euro-Mediterranean Disaster Information Network) "promotes the sharing of disaster-related information and data, research, results, knowledge and expertise. It aims at harmonizing methods to improve pre-disaster planning as well as hazard, vulnerability and risk assessments" (website: `http://www.eu-medin.org/`).
- GMES (Global Monitoring for Enviroment and Security) aims at an effort to "bring data and information providers together with users, so they can better understand each other and make environmental and security-related information available to the people who need it through enhanced or new services" (website: `http://www.gmes.info/`).
- ORCHESTRA (Open Architecture and Spatial Data Infrastructure for Risk Management) is another European Union project, which sets the goal to "design and implement an open service oriented architecture that will improve the interoperability among actors involved in Multi-Risk Management" (website: `http://www.eu-orchestra.org/`).
- OASIS (Open Advanced System for crisIS management) is an "Integrated Project which focuses on the Crisis Management part of improving risk management. It aims to define a generic crisis management system to support the response and rescue operations in case of large scale disasters" (website: `http://www.oasis-fp6.org/`).
- WIN (Wide Information Network) "targets at a wide range of opportunities at European and regional level, on several thematic, for instance, shoreline monitoring, water resources management, and various land-based risks management domain" (website: `http://www.win-eu.org/`).

The success of disaster management not only depends on well-defined policies and procedures, but also largely depends on the successful integration of related information to make decisions during the response phase. In the response phase, people involved have to deal with large amount of existing and operational data under a unpredictable and time-critical manner. Therefore the response phase is the most appealing one among all the phases of disaster management. During the response phase, it is needed to ensure the interoperability of emergency service (as well as information), to provide appropriate information at the right place and in the right moment [46].

This information ranges from existing data, such as topographic data, available transportation means, demographic data, to operational data, such as weather forecasts, available rescuing resources units distributions, measurements from the field and so on. Most of them are geographically related. Therefore when talking about integrating information for disaster management response, it refers to the integrating of geo-information. The importance of geographical information has also been mentioned by Goodchild F. M. [23] and by Cutter et al. [7].

However, successfully discovering and combining geo-information in a time-critical manner for use in decision making is not an easy task, because geo-information is distributed among different organisations. On one hand, those organisations possess parts of emergency related knowledge. On the other, they also collect, create and maintain data for other tasks and purposes. Exchanging those geo-information data (and knowledge) by interacting on a personal/phone/fax basis is a slow process, which may be prone to human error or interoperability issues. Developing methods to automate collection of geo-information, through machine automation, will enable the integration of needed information in a time critical manner to aid decision making during a disaster.

Current efforts to integrate geo-information data have been restricted to keyword-based-matching Spatial Data Infrastructure (SDI) [10]. SDIs are being set up within regions, countries or even across national borders ([4] and [33]) to facilitate the access to geo-information. SDI supports the discovery and retrieval of distributed geo-information sources and geo-information services by providing catalogue services and syntactic interoperability standards [22]. In the rest of this proposal, SII (Spatial Information Infrastructure) will be used instead of SDI.

An example of a recently started international SII is the European Union's INSPIRE (INfrastructure for SPatial InfoRmation in Europe) directive, which aims at making available relevant, harmonised and quality geographic information for the purpose of formulation, implementation, monitoring and evaluation of Community policy-making. It has the ideas that i). Data should be collected once and maintained at an appropriate level. ii). Seamless combination of spatial data from different sources and sharing of them should be enabled to various users and applications. iii). Spatial data should be collected at one level and shared among all different levels of government. vi). The discovery of spatial data should be made available ([31] and [20]). The US DHS (US Data model for Homeland Security) is the result from NSDI (National Spatial Data Infrastructure) in America [8]. The idea of NSDI, resembling that of INSPIRE, "is a physical, organisational and virtual network designed to enable the development and sharing of the nation's digital geographic information resources in the U.S.A." [8].

The integration of geo-information has been greatly advanced by SII. However, semantic interoperability, which challenges the integration of geo-information in the open and distributed environments, still exists. One possible way to deal with these problems is to make the formal semantics of geo-information available [20], so that different users involved in disaster management can exchange and integrate their data, with the help of machine automation,to do decision making in a time-critical manner.

Many efforts have been spent on specifying the semantics of geo-information nowadays. For instance, ICS (Information Communities and Semantics) Working Group at OGC (Open

Figure 1.1: Fire at a chemical plant and toxic gas leakage

Geospatial Consortium) "addresses the issues related to services that might be developed to help users manage semantic differences in spatial data content and metadata". W3C Geospatial Incubator Working Group aims at "addressing issues of location and geographical properties of resources for the Web of today and tomorrow ". It lies the groundwork for a more comprehensive geospatial ontology [21]. OpenKnowledge is an EU project aiming to "specify semantics in an appropriate way to provide a framework to enable a Semantic Web like system to become open, in the sense that anyone can join and the cost of individual participation is very low" [34].

Many projects that deal with semantic interoperability address the use of ontologies, for instance in [12], [28], [19]. In a recent PhD thesis, the author has also addressed the use of ontologies to describe the semantics of geo-services [20]. Agarwal, reviews the use of ontologies for GIS science and argues ontologies can help to deal with semantics interoperability problem [1]. Ontologies have been an emerging technology since the last decade. It is believed to be at least part of the solutions to semantic interoperability by describing formal semantics of geo-information. This PhD research with title "**Geo-information and formal semantics for disaster management**" focuses on using ontologies to describe formal semantics of geo-information for the domain of disaster management. The added value of using formal ontologies compared to traditional methods is going to be estimated.

Please note that the scenario and figures are taken from work done between 2002 and 2004 at the University of Muenster under the EU project ACE-GIS ((IST-2002-37724, http://www.ist-world.org/ProjectDetails.aspx?ProjectId=b0731d0ad18642f0a55c3cbc978ecd0b).

1.2 Motivation example — discovery and integration of geo-information

Consider the following scenario: In figure 1.1, an accident is reported to a fire station. An explosion has happened at a chemical plant; toxic gas is leaking from destroyed pipes. In order to evacuate the inhabitants of affected districts, a gasmal has to be created to plan an evacuation and guide the rescue team. The forecast of the gas dispersion is an essential part of this task.

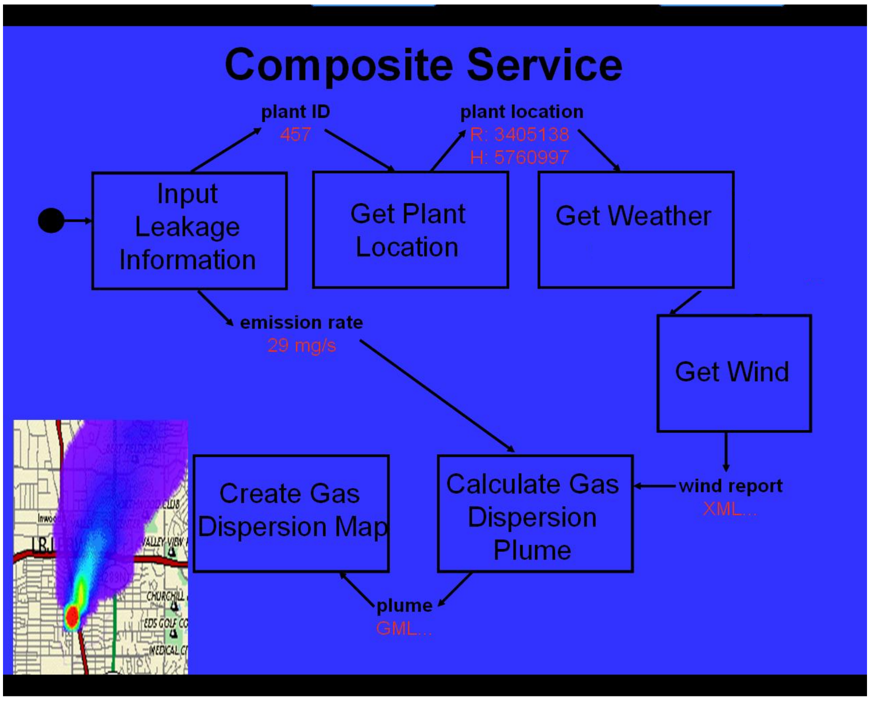


Figure 1.2: Integration from various data sources

In order to manage such a disaster, information from different sources needs to be obtained and combined immediately, which is shown in figure 1.2 (only an indication, not an exhausted list of all the needed information). For instance, the plant's ID, the plant location (the exact location of the emission), the weather information — the information of the wind from weather station (the wind speed and wind direction near the place of the accident), are depicted. Apart from what can be seen from figure 1.2, more information is needed, such as the domain-specific data (the population of the potential area, road network and so on), some on-the-fly measured data from the field (gas leakage), the weather forecast and so on, to create a gasmal to make the evacuation and rescue plan.

As long as the incident location is identified, the weather information around the incident area will be obtained. Together with the field measurements, such as the emission rate of the gas leak, these information is used to calculate the gasmal. The gas plume is presented in a map, with the gasmal on top of the road network and topographic data of the region (figure 1.2). In general, integrating different information and presenting in a single map supports the coordination of the evacuation efficiently.

Accessing and processing these heterogeneous information sources shall be done in the form of a service. The service arranges multiple web services within a service composition. Such a composition will be produced by a service composer whereas the gasmal is presented in the resulting map. The processes of information discovery as well as integration are essential within this scenario.

However it remains a problem as how to employ machines to help automate (or semi-automate) the process of discovery and integration of information because the discovery and integration of information in disaster management is time-critical and at the same time it is impossible (or very difficult) for humans to do such when facing large amount of information.

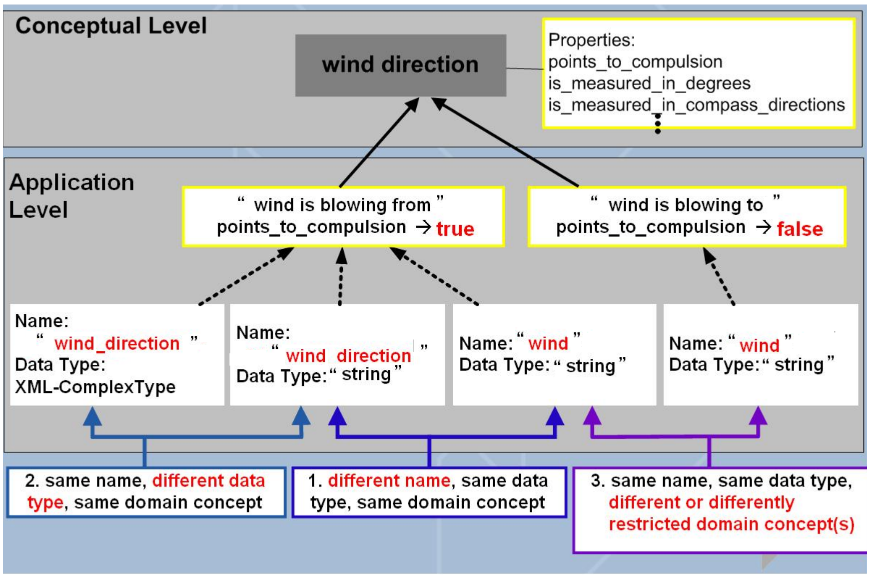


Figure 1.3: Semantic Interoperability

As shown in figure 1.2, it is required to find the information of the wind direction and the wind speed("Get Wind" in figure 1.2), which can be achieved by typing and searching the keyword "wind" or "wind direction ". A large amount of results will be returned, among which will be chosen for answering the information about the wind speed and wind direction. But there exists some problems which are illustrated in figure 1.3:

- The attribute name "wind direction" in different geo-services (or data sets) may mean the same thing, but with different data-type for storage (e.g., one with XML-Complex Type and the other with String Type). (Same name with the same domain concept but with different data-type.)
- The attribute names " wind direction" and "wind" in different geo-services (or data sets) are referring to the same thing— the direction of the wind blowing from. (Different names with the same domain concept and with the same data-type.)
- The attribute name "wind" in different geo-services are referring to different things— one refers to the direction which the wind is blowing from and the other refers to the one which the wind is blowing to. (Same name with same data-type but with different domain concept.)

One needs to figure out what "wind direction" means in each data sources by examining these heterogeneous data sources separately. However this trivial example is only one step among many in disaster management. Without ontologies, too much time and efforts will be wasted on examining heterogeneous data sources. By using ontologies, the semantics of these heterogeneous data sources are made explicit. Rather than examining the heterogeneous semantics for these data sources, the semantics of these data sources are made available by ontologies, which are able to be processed by machines.

Ontologies are needed to formally describe the semantics of information so that machines can help with the discovery and integration of information. And consequently, the disaster can be managed in a much more efficient way. Under this idea, the proposal advcates ontologies. With the help of ontologies, the meaning (semantic) of the information should have been made explicit and machine-processable. In this PhD project, ontology is proposed to be used to support information discovery and integration in disaster management.

1.3 Interoperability problems — Different hierarchical levels and problems with integration of geo-information

The management of a disaster is a complex process. Due to various types and scales of disasters, multiple organisations participate in the management work and numerous data are required. It is also very unlikely that all agencies will ever standardise on a single system. In reality, even within individual agencies there are multiple systems. This diversification is true for disaster management. This section takes the management of disasters in the Netherlands as an example to illustrate the complexity of disaster management and interoperability problems during disaster management.

1.3.1 Different hierarchical levels

Disaster management response involves multiple levels of cooperation of users from various organisations, as well as different scales (different scales of details) of information that are required by multiple users. For example, in the Netherlands, the Fire Brigade, the Police, the Medical Service and the Municipality are the first responders. Each of them is maintaining their own data and is carrying their own daily work. In case of a disaster, they have to cooperate with each other to deal with the disaster. Other institutions and organisations such as Red Cross, Military forces, Ministry, and more extra-specialised organisations may also get involved in managing the disaster when the magnitude of the disaster increases. They require information at different scales of details. For instance, the Fire Brigade needs more detailed information (such as the inner structure of a building) than the Queen does (the area information).

To specify the cooperation and communication between different parties, 25 processes and six GRIP (Common Regional Incident management Procedures) levels have been defined [45]. The disaster management is carried out as a combined action of several processes. Processes are indicated as a series of connected activities, which take place simultaneously depending on the type of the disaster. In the Netherlands, 25 processes are defined and organised by clusters. For each cluster, there is one responsible sector (the Fire Brigade, the Police, the Medical Service and the Municipality). Process 05 (observations and measurements) will be activated in the example in 1.2, which will measure the leak of the toxic gas.

Apart from the processes, GRIP levels are also defined regarding to the scale of the disaster. On GRIP 0, Fire Brigade, Police Medical Service and Municipality are doing their daily work. There is not too much cooperation work among them. However, during GRIP 1, the mayor of the Municipality will be informed. A cooperation team will be formed by the representatives from each of these organisations. With the increase of the magnitude of the disaster, more organisations are getting involved. For instance, emergency officers at provincial or national level are informed if the disaster affects a large section of the community; the Ministry of International Affairs will take the administrative lead if the disaster extends beyond the provincial; secretary of the the Queen will be informed if the disaster crosses the national security (e.g., a nuclear leak). So with different levels of GRIP, different types of cooperation are established and strict policies are used to form the structure to deal with disasters.

At the same time, users at different hierarchical levels are requiring different details of information during disaster management. People working at the field are requiring more detailed information than those who work at higher levels. For instance, the Fire Brigade uses maps that contain access paths for each building while the mayor of the Municipality uses GBKN or TOP10NL maps, which are more general. Even the same user may require different levels of details or different content of information. For instance, a fire brigade will participate in one's daily work, such as giving advice to the storage of dangerous goods, or checking the status of the buildings while in disaster management one is involved in giving the lead in the field, reporting the status of the field and etc.

1.3.2 Integration of geo-information

As discussed above, different data sources are requested and have to be discovered, accessed, collected and integrated. These data range from existing(rather static) data, such as topographical data, to operational (highly dynamic) data, such as the measurements from the field.

1. Existing data are those that are created and maintained by organisations before the disaster happens. The existing data includes general purpose data and specific data. By general purpose data, it refers to, i.e. the population data, hydrological data, ground water data, the pipeline data, the topographic data and so on.

   The specific data refer to those that are maintained by individual partners, especially the municipalities and provinces, which have a multitude of these data available. Specific data are, i.e. transportation data, weather forecast and so on.

2. Operational data are those received from the field during disaster management. They are not available prior to a disaster. They could be a report (description) of an incident, camera images, video clips, measurements, aerial surveys and so on.

Various users from multiple organisations have their own specific views on the problem, they need different sub-sets of information from desperate data sources or combination of different data sources. But these data are maintained or created by various organisations, for example, Police, Fire Brigade, Medical Service, Municipalities, and some specialized organisations (such as the mapping agencies, utility companies and so on), thus they are heterogeneous in data model creation, data maintenance and data presentation:

- Data model creation: Data are used for different purposes. Users create their specific data to meet their own needs. They look at the problem from their own point of view. Even when they are referring to the same problem, they might have different interpretations. For instance, talking about water, users have distinct interests in it. The people working for forecasting floods might only be concerned the height of the water, while the people working for measuring the pollution might only have interests in the chemicals in the water. Thus this results in different ways of interpreting and modeling of the real world.

- Maintenance: Data are stored in numerous formats. organisations are using various technologies to collect and create their own data. Different storage systems are employed to update and maintain the data. For example, the data could be files, a relational database, graphics, text, picture or video clip and so on.

- Presentation: Users are used to work with particular representations of data. And the presentations are sometimes combined in data models. For instance, colors are usually associated with meanings. The fire brigade associate a red color with danger and a green color with safety, while road surveyors use the colors only to distinguish among different road networks.

Interoperability problems will arise when integrating these individual data sources. When talking about the word “interoperability”, it means that “the ability of systems to provide services to and accept services from other systems and to use the services so exchanged to enable them to operate effectively together” [27]. The idea of interoperability is due to the fact that heterogeneity exists in information system, so it is the same with disaster management.

Due to different types of heterogeneity, there are corresponding different types of interoperability. Normally, the division is made as system interoperability, syntax and structure interoperability and semantic interoperability [35]. System interoperability refers to the ability to deal with hardware, operating systems, and communications heterogeneity, such as the instruction sets, communication protocols, file systems, naming, file types, operation and so on. Syntax and structure (schematic) interoperability is relevant with machine-readable aspects of data representation, formatting, data models, different DBMS (Data Base Management System). Semantic interoperability has more to do with the meaning of the data content.

Semantics is the study of the meaning of expressions. It refers to the aspects of meanings that expressed in a language, code, message or any other form of representation. Sheth argues that semantic interoperability requires that the information system understands the semantics of the users information request and those of information sources, and uses mediation or information brokering to satisfy the information request as well as it can”[35].

The semantic interoperability problem is not evadible due to the fact that the use of different terms and approaches cause confusion in the specification of universally accepted entities, concepts, rules, and relations. Apart from that, the semantics of the creation, maintenance and representation of geo-information are not clear to the people outside the organisation. Thus the semantics of these related data for disaster management response should be understood by all the users involved in disaster management response, so that they could exchange their information under high time pressure.

In order to achieve that, it is necessary to make the semantics of geo-information available not only for humans but also for machines— the semantics of geo-information needs to be described in a machine processable way, so that the process of exchanging and integrating large amount of data could be automated.

Chapter 2: **State-of-the-art technologies**

2.1 Formal Semantics

Formal refers to the fact that something is represented in an artificial and well-defined language so that machines can process it. Usually four degrees of formalities are used in the real world. They are, namely, *informal, semi-formal, formal and rigorous formal.* Informal means knowledge is represented using human language. Semi-formal refers to the fact that knowledge is expressed in a restricted and structured form of natural language, for instance apply some patterns with human language. Knowledge represented at these two levels is aimed at exchanging knowledge between human beings. Formal and rigorous formal represented knowledge is designed for machine processing. They are both defined in an artificial and well-defined language, while the latter one includes theories and proofs of properties such as soundness and completeness.

Semantics, as discussed above, is the meaning of the information, message, words, maps (graphic and symbols) and so on, which is expressed in a language, programmes, or other forms for representation. Formal semantics specifies the meaning of the information, message and so on are represented in a well-defined and artificial way so that it could be processed by machines.

If the formal semantics of geo-information can be successfully described, it will benefit from three aspects, which are shown in figure 2.1:

1. Data content: This concerns the content of the information — locate and select relevant information that satisfy the service request. For instance, "give me all the relevant information that is about the orientation of the road".
2. Service chain: This refers to the integration of the information — composite and mediate relevant information based on their content (semantic). Based on the content (semantic) of information, combine different information sources and solve mismatches among them. For instance, the software which computes the gasmal needs to integrate the information from the field measurement, the weather information and other relevant information.
3. Machine interface: This is regarded as, get the result and present it to the user — an automation of a question-answer process. For instance, in case of a fire at a chemical, input the relevant disaster type and alarm information, get a suggested evacuation plan ("Input: fire A happens at location B, output: evacuation plan").

2.2 Ontologies

Ontologies is one of the emerging technologies that are used to describe formal semantics in the past 10 years. In much research, ontology is used to represent the formal semantics. In the following sections, ontologies are to be discussed.

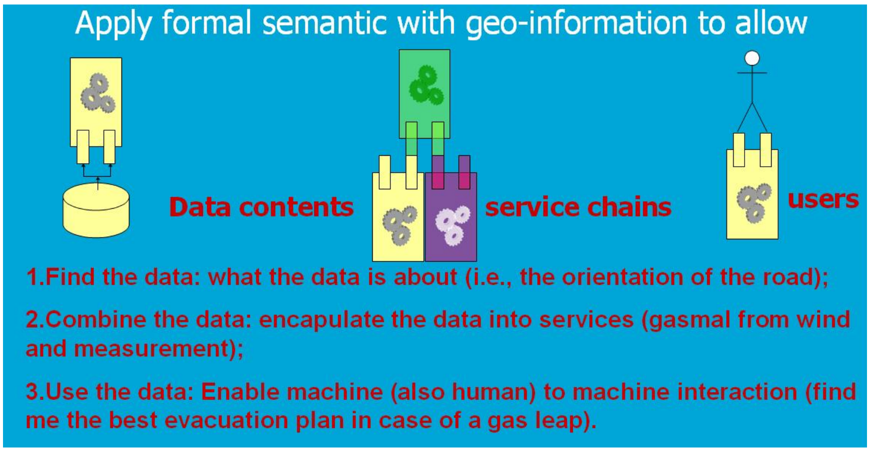


Figure 2.1: Benefits from describing the semantics of geo-informatiion

The word ontology, originates in philosophy, which is "the study of being or existence". "It seeks to describe or posit the basic categories and relationships of being or existence to define entities and types of entities within its framework" [14].

Computer science borrows this word from philosophy and according to Neches, one of the earliest definitions for ontology in knowledge engineering, gives the definition of ontology as "defines the basic terms and relations comprising the vocabulary of a topic area as well as the rules for combining terms and relations to define extensions to the vocabulary" [29].

Gruber defines ontology as "an ontology is an explicit specification of a conceptualization"[13]. Further, he proposes a terminological clarification among different uses of ontologies as, an "informal conceptual system", a "formal semantic account", a "specification of a conceptualization", a "representation of a conceptual system via a logical theory", the "vocabulary used by a logical theory" and a "specification of a logical theory"[12]. Swartout in his paper argues: "An ontology is a hierarchically structured set of terms for describing a domain that can be used as a skeletal foundation for a knowledge base"[42]. Borst refines and extends the definition of ontology as "ontologies are defined as a formal specification of a shared conceptualization"[5].

These definitions are not exactly the same but they all reveal two key points of ontology i) ontology is a kind of conceptualisation and ii) ontologies should be shared. According to Studer and Uschold, conceptualisation is an abstract model of the real world phenomenon, by which the real world is identified as a set of concepts. Shared means the notions are accepted by a certain group as consensual knowledge [41] and [44]. In the following two important aspects in defining different types of ontologies will be highlighted:

- Explicit. Explicit means the meanings of the types of concepts that are used in the conceptualisation and the constraints of their usage are defined [41] and [44].
- Formal: Formal refers to the fact that the ontology is defined in an artificial and well defined language so that it is machine-readable [44]. Ontologies could be defined in other way with respect to the degrees of being formal. An ontology could be defined informal if it is written in natural language; semi-formal if it is expressed in a restricted and structured form of natural language; formal, as it has been described; and rigorously formal if they are defined in a language with formal semantics, theories and proofs of properties such as soundness and completeness [44].

In a word, ontology could be defined explicit or non-explicit (implicit) and with different degrees of being formal and being shared. As adopted from Audi, formal explicit ontology is defined as "the study of explaining reality by breaking it down into concepts, relations and rules", and then describe them formally and explicitly and share it with others. Usually an ontology is composed of four components [3]:

1. Vocabulary (controlled vocabulary): "a list of terms that have been enumerated explicitly" [16]. All terms in the controlled vocabulary have an unambiguous, non-redundant definition for a domain. For instance, the words "building", "storage", "function" are defined as the vocabulary for a certain domain, which means these terms are understood and shared within this domain. There is no need to further define these terms.

2. Classes (concept): a set of objects and the basis of knowledge representation in an ontology. It can represent every entity in the real world (a task, a function, a process, an object). For instance, the class "Warehouse" is defined as "the building whose function is for storage."

3. Relations: represents types of interactions (is_a, part_of, lead_by) between the classes (concepts) in an ontology. For instance, the class "Warehouse" is related with the class "Goods" by the relation "stores", which means "The warehouse stores goods".

4. Rules (axioms, or constraints): conditions that is always true for a domain. For instance, the fact that "a warehouse can be used as a shelter" can be expressed in a rule "the function of a warehouse equals to the function of a shelter".

The distinction between Ontology modeling and OM (Object Modelling) is not that obvious. And they could be replaced with each other to some extend. Since Ontology modeling originates from knowledge representation and the object modeling has the root in information engineering community [11], Ontology modeling are more concerned with "what qualifies a being in the real world", while object modeling is more focused on "what exists in the real world".

It is necessary to point out that, in order to solve the semantic interoperability and allow machine-automation of data integration, it is desired that the semantics of the data be defined explicitly and represented in a machine-process-able way. Ontology can help model the domain of interest. Explicit and formal ontolgies can help define the semantics of the data, make them sharable by different users and allow machine-processable. So when talking about ontologies in the following sections, it is referring to explicit and formal ontologies.

2.3 Ontologies architectures

Three different approaches for using ontologies have been identified and a division has been made into , *a Single (global) ontology, Peer-to-peer (multiple) ontologies* and *Hybrid ontologies* in the following manner [48], [49] and [43]:

- In a Single ontology, one global ontology is used to provide a shared vocabulary for specifying the semantics. All information resources have to use this shared vocabulary or a subset of it. As a consequence, all information sources have to share the basic understanding of the given domain. Relations among different information sources can be inferred automatically.

Within this model, every domain concepts and relations are encoded. It provides a clear and complete overview over this domain. It is possible to have such one, only when all the different parties involved in the domain of interest have a common agreement over the domain and no conflicts within the domain. Otherwise it is hard to harmonize such an ontology model.

Figure 2.2 a illustrates the idea. The SIMS project [2] is an example of this model, where a global called SIMS model is built which includes a hierarchical terminological knowledge base with nodes representing objects, actions and states. It also includes indications of all relationships between the nodes. In the system, each resource is simply related to the global domain ontology.

- Peer-to-peer (multiple) ontologies are characterized by demanding a distinct ontology for each service. These ontologies do not have a common base, so it is difficult to compare two services to each other. The only way of solving this task is to explicitly formalize cross-ontology mappings. Following this approach, all transformation rules from one ontology to another need to be defined manually. As a result, the relations between two services can be identified by considering the transformation rules between their related ontologies.

  Peer-to-peer ontologies make sure this is a mapping between any of the parties. And it is easy for the domain ontology developer to build such ontologies, because the developer only needs to consider the mapping between two parties. But the drawback is it results in large amount of mappings — $n^2 - 2n$ (n is the number of the parties).

  Figure 2.2 b illustrates this idea. Uitermark uses a peer-to-peer ontology to integrate two topographic data sets, GBKN (large-scale topographic data set) and TOP10 Vector (medium-scale topographic data set) [43].

- Hybrid ontologies are developed to overcome the drawbacks of the previous two approaches (figure 2.2 c). Each service is referenced to a local ontology. But in order to make them comparable, a shared vocabulary is used to build the local ontologies. The shared vocabulary contains basic terms (primitives) of a domain.

  The advantage of a hybrid ontology is that new sources can easily be added without the need of modification. And it reduces the number of mappings from n^2 to n. The hybrid ontology is adopted, since it more suits the situation of disaster management—different data are to integrated and different people from different hierarchical levels of organisations are involved.

2.4 Approaches to create ontologies

Generally speaking, there are two approaches to build ontologies, namely, top-down approach and bottom-up approach [1]. They both have their own pros and cons. The two approaches will be compared in order to find an approach of building the ontologies.

- Top down ontologies development indicates an ontology is constructed by first examining the domain of interest in general at a very abstract level and then refine our constructing based on top levels concepts. This is accomplished by building an abstract model of the domain of interest first (or by starting from existing upper level ontologies, which are explained in next paragraph) and then extending the model further to map more specific concepts from low levels. For instance Kassel uses DOCLE [25] as a starting point for elaborating new ontologies [18]. DOLCE is the Descriptive

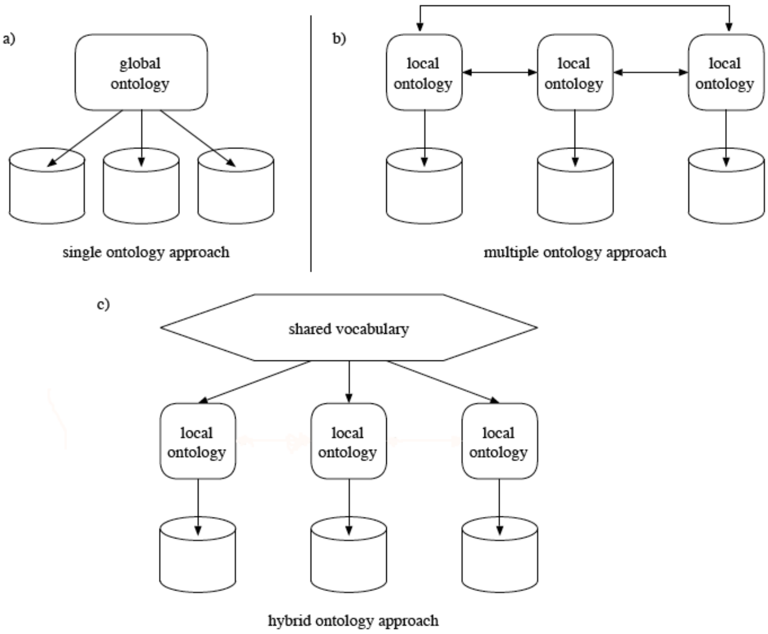


Figure 2.2: Different ontologies. The arrows indicate mapping rules among different ontologies. Figure adapted from [49]

Ontology for Linguistic and Cognitive Engineering, aims at capturing the ontological categories underlying natural language and human commonsense.

In Agrwal's report, a upper ontology, a top-level ontology is an attempt to create an ontology which describes very general concepts that are the same across all domains [1]. Apart from DOCLE, there are many other top-level ontologies, for example SUMO ontology (Suggested Upper Merged Ontology, a foundation ontology for a variety of computer information processing systems) [30], CYC (a comprehensive ontology and database of everyday common sense knowledge, with the goal of enabling AI applications to perform human-like reasoning) [40], Basic Formal Ontology (a ontology about theory of part and hole, theory of dependence and theory of boundary, continuity and contact) [36] and [37], and so on.

There is research carried out with respect to the usages of top-level ontologies. Probst provides formal semantics for the central OM (a conceptual model and encoding for observation and measurement) [6] terms by aligning them to the foundational ontology DOLCE [32]. The alignment to a foundational ontology restricts the possible interpretations of the central elements in the OM model and establishes explict relations between categories of real world entities and classes of information objects. Stubkjaer talks about the usability of CYC ontology for building cadastral application ontologies [40]. Simth and Medo illustrates the usage of using BFO to build ontologies for bioinformatics and medicine [38] and [37]. How to adopt a foundation ontology based on which to extend remains a question.

The resulting ontology from top down approach is a complete model over the domain of interest. It contains all the relations between the concepts within the domain of interest. It is also useful for other domain ontology developers, who can build their ontologies based on this one (ontology extension in ontology construction), or map their ontologies (ontology comparison in ontology usage) with this one.

When building such an ontology, there are some challenging issues to notice. Ontology developers and domain experts need to cooperate and explore over the domain of interest and then discover core concepts and relationships within this domain. However it remains a problem for domain experts to reach an agreement over a domain, because it cannot be assumed that people look at the problem in one single way — they might have different tasks, rules(specification) and views over one problem. Second, it is difficult to extent the top-level concepts with lower level data sources at the same time maintaining the model integrity. Sometimes, the top-level structure through top-town approach might be wrong and can not be extended to detailed levels. Building such an ontology requires the ontology developers not only to have a thorough understanding of the domain of interest in depth but also to have a clear overlook over the domain of interest at wide breath.

- Bottom up: Bottom up approach begins by looking at existing information sources from low levels (data schema, data structure labels and etc.), developing ontologies for specific individuals, and then combining them as a whole. The process starts from the low-level information sources and moves towards a higher level of abstraction. The result of this approach is an accurate model carrying mappings among all the individual data sources with the domain of interest. For instance, the process of building an (information exchange) core model for a domain usually takes this approach. A core model, here we mean, a set of object model concepts that all domain members must support.

  When building such an ontology, too much attention is being paid to the details of the information, such as the structures of the information, the implementation details of the information. Many compromises are being made for information harmonisation, such as the inconsistency, overlapping and presentation of the information. This ontology could not be a complete model to represent the relationship of the domain of interest because it is too focused on the information sources specifications.

  Bottom-up approach is usually used to refine ontologies, because sometimes core models of different domains may not be consistent. We need to refine the ontology created from the top-down approach using bottom-up approach. For example, a core model for cadastral model (which is created using top-down approach and defined by bottom-up approach, with respect to the cadastral domain) is given and a translator has been made between this core model and two other different cadastral datasets [47] and [15].

  On the other hand, since it focuses on the specifications of the data sources, such an ontology captures the relationship between the data sources well, such as the inconsistencies, overlapping, disjointness, equivalence and so on. It works well for data exchange.

  The approach of building such an ontology does not require ontology developers to have as thorough understanding of the domain of interest as the top down approach demands. But instead it requires the ontology developers to be aware of the differences among different data sources. It is easier for ontology developer to start with by looking at the data sources first, because it does not require the ontology developers to have a thorough overlook at the whole domain.

2.5 Ontologies for Disaster Management

Due to the characteristics of disaster management that different hierarchical levels of organisations are cooperating, large amount of data from different places need to be integrated, as well as different scales of details for information are required, we will adopt hybrid architecture ontologies. In other words, desperate information will keep their own local ontologies and a "glue" is going to be built that connects them without changing them.

In the following, some terms are introduced, such as 'data ontologies", "ontology for processes", "local ontologies", "domain ontologies". These terms reflect the fact that that there are different ways to categorise ontologies. When "local ontologies" are used, it means this is a particular ontology for one data source. For instance, a local ontology for TOP10NL data set (a topographic data set), a local ontology for cadastral data set (a cadastral data set). "Data ontologies" and "ontology for processes" are terms that are used to describe the usage of the ontology. "Data ontologies" are used for describing the data while "ontology for processes" are used to describe the processes in disaster management (the detailed descriptions of the two can be found in the following paragraphs). "Domain ontologies" refers to the whole set of ontologies that are developed for disaster management. It merges the data ontologies and ontology for processes and encapsulate them as one.

Ontology for processes is the one that describes the processes in disaster management, which can be considered as the work flows for disaster management — how the disaster management is organised, what are the responsibilities of each user, how the users communicate with each other, which user participates in which process and which process needs what information. We will examine through the processes (there 25 to 29 processes in the Netherlands) and GRIP (GRIP 0 to GRIP 2) levels to work out what information is required by which user.

Data ontologies are a aggregation term, which consists of several ontologies that describe the information sources seperately (e.g., topographic data sets, utility data sets, cadastre data sets and so on) that are needed for disaster management. They are independent of the domain of disaster management, because all these data sets could be used for other domains, such as environmental domain, and water management domain. Data ontologies form the foundation for information integration. In disaster management, there are plenty of data not only from the existing databases (e.g., the plant information in figure 1.2) but also from the field (e.g., the measurement of the wind in figure 1.2) that need to be processed and combined (e.g., the gasmal in figure 1.2). For each of these data, a local data ontology will be built describing the content (and structure) of the data.

Generally speaking, ontologies for processes serve as the glue that integrates different data sets together. It helps the disaster management to discover and combine desperate data sets. Data ontologies specify what is included in the data source and how the relevant information can be achieved. The ontology for processes together with data ontologies form the ontology for disaster management.

Chapter 3: **PhD research**

3.1 Research Problems

In order to make it short and obvious, it summarises the above sections and define the main questions of the PhD research as:

What is the added value of formal semantics for geo-information, compared with current solutions (with knowledge hidden in the "hard-codes" or some models)? How to apply formal semantics (in this case ontologies) in disaster management, so that some man-made decision-making can be maximally supported by machines?

Under these research questions, many sub research questions can be defined.

1. Are there current tools sufficient to describe the formal semantics of geo-information for disaster management? Why it is needed to employ ontologies to describe the formal semantics? Are current tools sufficient to do data integration?

2. Are ontology languages more powerful than the commonly used modeling language, such as UML. Since many of the previous works are made using UML, how to translate the models in UML into one using ontologies languages, for instance, OWL (Web Ontology Language) and RDF (Resources Descriptions Framework)? Are there any tools which translate ontology models from one form to another (i.e., from UML to OWL)? For instance, ODM (Ontology Definition Metamodel), which is "an emerging standard from the Object Management Group that supports ontology development and conceptual modeling in several stand representation languages", an appropriate framework for "bridging between the knowledge representation (such as Ontology representation) and the information system engineering communities (such as UML representation)" [11]?

3. Is the current ontology modeling language sufficient to model the decision-making processes? What are the cons and pros of the existing modeling (ontology) language, for instance, UML (plus OCL), RDF [24], the three series of OWL— OWL Full, OWL DL and OWL Lite [26], and so on. Is it necessary to include some logical programming or logic languages, for instance, first order logic, to help model the processes of decision making?

4. Could be the represented knowledge interpreted to some logics, for instance Description logics, so that some existing reasoners can be used to do reasoning with the represented knowledge. Is it necessary to develop a specific reasoner for specific use or develop a more general-purpose reasoner with more functionality.

5. How to start building such an ontology? — either from the abstract level, from concrete level or approach from both sides? The proposal aims at developing ontologies that are both good for exchanging information and suitable exchanging domain knowledge.

6. Among different ontology architectures, such as hybrid architecture, peer to peer architecture, global architecture, select a suitable one for building the ontology for disaster management.

7. How to store the represented knowledge? Is it necessary to employ DBMS (database management system), such as Oracle, for managing the represented knowledge? What are the pros and cons if DBMS is employed to manage ontologies?

8. How to query the represented knowledge? If they are stored in a DBMS, is it possible to do query using SQL or other query language, such as SPARQL.

3.2 Relations and roles with projects involved

The PhD research work is carried out between two projects, which provide relevant experiences with the domain of disaster management and geo-information:

- GDi4DM, Geographic Data infrastructure for Disaster Management, which aims at the development of a Geographic Data Infrastructure to support decision-making and information exchange during the disaster management phase to facilitate the interoperability of emergency services and the quick and effective exchange of accurate information.
- Humboldt, Development of a Framework for Data Harmonization and Service Integration. This project will contribute to the implementation of an European Spatial Data Infrastructure (ESDI) that integrates all the diversity of spatial data available from the multitude of European organisations.

The project GDi4DM and Humboldt are two use cases for the formal semantics of geo-information. The former one aims at the interoperability of emergency services and quick and effective exchange of accurate information for disaster management. The PhD research will primarily based on the work of GDi4DM to examine the relations among different and heterogeneous datasets, the complex communications among different users involved in disaster management and the relations between the users and data. The result will be ontologies for GDi4DM to support decision-making and automatic information exchange in the Netherlands.

Other projects, for instance, W3C geospatial Incubator [21], OpenKnowledge project [34], Common Alert Protocol [17], FGDC Emergency Symbology, and European Mediterranean Disaster Information Network (which has been mentioned in section 1.1) will closely be followed. Some developments in practice will also be examined as an input for ontology creation, for instance, US DHS [8], EU INSPIRE [31], FIG/ISO CCDM [47] and so on. All the mentioned projects and models have relevance with respect to the PhD research to some extend — either at technical aspects and at the strategic aspects.

3.3 Research Methodology

The PhD research involves the work from theories to practice and back to theories again. So the research will include works from theories studies to practice and back to theories research and so forth. The following gives a summary of the methodology, which will be carried during the PhD research.

3.3.1 Literature Study

In order to avoid repetitive research work and get inspired by previous works, reading literatures is not evitable for the PhD research. The works includes collecting related background information for disaster management, collecting previous knowledge and works that have been done to deal with disaster management, and understanding the normal procedures in case of a disaster management. At the same time, from a technological point of view, collecting knowledge that are related with technologies is also important.

Emphasis of this phase will be on semantic web technology, logic programming, modeling theory, database technology together with policies and procedures of disaster management.

3.3.2 User Requirements Exploration

Since every technology should facilitate a group of users, the PhD research work should also understand the target users — people involved in disaster management. By understanding the inner relations between different users and related datasets, it becomes clear what is going on within the domain of disaster management.

The success of a technology product comes from an iterative cycle between understanding the user requirements and the development of the product [39]. The PhD research will be based on users requirement exploration. The user requirements that are collected within GDi4DM will be used primarily.

3.3.3 Data used Examination

One of the PhD research works includes studying the data for disaster management. As it has already been discussed that these data range from existing data maintained by different organisations to operational data from the field of a disaster.

With existing data, it refers to the general purpose data and specific data. For instance, overview maps of Netherlands (overzichtskaart Nederland) with the scale of 1:1,000,000, topographic maps (Top-10 Vector, Top-25 Raster and Top-250 Raster), Cadastral maps, air photography, GBKN (large-scale maps of Netherlands with the scale of 1:1,000), current altitude data of Netherlands (Actueel Hoogtebestand Nederland), are all general purpose data. Cabel and piping, statistic data agrarian comit, plot and plant information, roads, water ways, track ways, station allocation, ports, shipping routes, shipping movements, airports, flight routes, dangerous-substance allocation and so on are examples of specific data.

With respect to operational data, the digital forms of them are not available, such as dynamic data coming from the measurements, reports and etc. from the field. Some of these are not available at the moment, but they will be obtained from the work of GDi4DM.

3.3.4 Developments— Building Ontologies

As explained in previous section, the *hybrid ontologies architecture* will be adopted, which will result in several ontologies. The data that are needed for disaster management can exist with or without (these data are referred to as existing data, section 1.3.2) the context of disaster management. For instance, the road maps can be used for route planning during disaster management as well as for urban designing; the basic registry database that is maintained by the municipalities can be used to manage population during the disaster managements. Thus it helps allow more interoperability by building ontologies for the data, which is independent on the context. Bearing this idea in mind, the data that are needed for disaster management, e.g. GBKN, TOP10NL, basic registry database and so on, will be examined and ontologies, which are independent on the context, will be developed separately.

The ontology for processes will be developed according to the users' requirement for the disaster management. The users' requirement describes the work flow of disaster management — which actions should be activated if an incident happens, which actors should get involved, which information should be collected, what are the relations among different emergency officers, etc. During the disaster management, a large amount of data (these data are referred to as operational data, section 1.3.2) will be created as well, e.g. the measurement of air pollute, the report of victims. These data are dependent on the context of disaster management. Thus the when the development of ontology for processes finishes, the ontology for the operational data which are created during disaster management will be developed too.

Having the ontology for processes and data (existing and operational) ontologies ready, an ontology for disaster management will be developed. This is one of the final goals for this PhD research (as well as testing the usage of the ontology). The ontologies will integrate the ontology for processes and data ontologies, which describes the terms, concepts (users, geographic objects) and relations (work flows, data) for disaster management.

3.3.5 Developments— Building a prototype

The research problem originates from real world problems and we are seeking a solution for that. That's why tests will also be essential in this PhD research.

The hypothesis of the result of PhD research will be a system, which allows the machine automation of the discovery and integration of geo-information for disaster management. Formal semantics of geo-information for disaster management will be described and a series of ontologies (or shemas) will be developed for the management work of disaster management. The final demo will consist of several components — web services, ontologies, a reasoner, a user interface, a virtual database which consists of several real databases. A user interface is provided with the user to input their request and output the result of their request. The reasoner together with ontologies connects several database as a virtual database and presents the result to the user. So it seems to the user that, he/she is communicating with one instead of several databases.

3.3.6 Cooperation

The PhD research is such one that covers many other research fields, for instance, knowledge engineering, geomatics, disaster management and so on. So collaborations will be established with related research organisations, companies and persons. For instance, a cooperation with the University of Free Amsterdam.

3.4 Topics beyond the PhD research

As the topics of the PhD research covers many research areas, there are many problems existing or related with the PhD research. It will point out some of them in the following but they are beyond the PhD research.

- Distributed environments. In this PhD research, it is assumed that the data is stored in one place and full access to all the data are granted to the users. However in real situation, the data are distributed at different places and users do not have the same access to the data. In this PhD project, it only concerns which user get what data through different combinations of processes and GRIP (only GRIP level 1 and GRIP level 2 will be considered).
- 3D model. Even through the 3D model is a hot topics and 3D model data are useful for disaster management, it is not going to be used for the PhD research.

Chapter 4: **Research Plan**

Since the PhD research is related to the project (GDi4DM and Humboldt) development progress and several other activities (for instance, the education), the following schedule is only a rough one, which indicates the time consumption on the research work of PhD. The PhD research officially started on September 1st, 2006.

As other common PhD plans, 70% of the total time will be made available and spent on the PhD research work, 5% of the total time will be spent on writing paper and reports, and the other 25% will be used for other issues, for instance, attending educations, conferences, meetings, other research projects, supervising students and so on. The overall plan is as following:

- The last four months (September, October, November and December) in 2006 and the first month (January) in 2007 will be spent on background literature studies and the writing the PhD proposal.
- Starting from February 2007, untill 2009 (also the first one or two months) will be the main period for conducting the PhD research (3 years in total). See figure 4.1 and section 4.1 as details.
- 2007 (1 months) Writing research plan (part 2) 1 month.
- 2010 The last six months will be spent on the writing of doctoral dissertation and preparation of defending ceremony.
- During the research period, a plan for visiting other universities or research institutes is considered. The total amount of time will not exceed 4 months.

At the same time, this is a publication goal together with the PhD research:

- Each year a progress report will be made for GDI4DM (and or for Humbolt).
- Each year two conference papers are planned for submissions regarding disaster management and semantic web conferences.

4.1 PhD research agenda

1. Preparation.
    - Literature Study. This phase is the preparation phase for the PhD research, which includes: i) understanding the background of the PhD research — disaster management; ii) reading and comparing of past or on-going projects related with disaster management and so on. This takes about four months — from September, 2006 to December, 2006.

- Writing of PhD proposal. The whole month of January and February in 2007 will be spent on the writing of PhD proposal. Starting by writing a draft version and get feedback from the supervisors. The result will be an official PhD proposal and published as a section report.
- Software preparation. The whole month of April in 2007 will be spent on collecting technical information and selecting proper softwares for the following work, for instance, semantic web technology, ontology technology, tools for developing ontologies and so on.

2. Local data ontologies building.

   - Local ontologies building for ED (existing data). Study related existing geo-information data needed for disaster management response, examine detailed descriptions (or schema) of different data-sets, and build local ontologies for these data sets. The result will be helpful for finding their relations, such as overlapping, dependency, inconsistencies and so on. The result will be local ontologies for each data sets.
   - Local ontologies building for OD (operational data). Study related operational data needed for disaster management response. Understand their roles in the management work of disaster response. Build local ontologies for operational data will be helpful in finding their relations, such as overlapping, dependency, inconsistencies and so on. The result will be local ontologies for the operational data for disaster management.

3. Ontologies (Vocabularies) building.

   - Ontologies building for ED (existing data). When the existing data (local ontologies) are ready for integration, we will building ontologies for the existing data so that they are comparable with each other. The ontologies serve as the "glue" that combines existing data sets together. The result will be ontologies (or shared vocabularies) for the existing data.
   - Ontologies building for ExOp data (Existing data and Operational data). In order to combine the existing data and operational data together for disaster management, we need a "glue" to connect them together. That "glue" is the ontologies for existing data and operational data. The result will be ontologies (shared vocabularies) for existing data and operational data.
   - Build the ontology for processes in DM (disaster management). The targeted benefactors of the PhD research are those who involved in disaster management. Understanding their requirements plays an vital role in the PhD research. This step includes visits to different departments, for instance, the Fire Brigade, the Police, the Municipality, some other mapping agencies and so on. Understand their main role and their relations when they cooperate to deal with disasters. By examining the processes in disaster management, the inner-relations and inter-relations among different users from different organisations will be examined. And the result will be an ontology for processes in disaster management.
   - Domain ontologies building. When both the ontology for processes has been developed and all the data (integration of existing and operational data) need for disaster management have been made ready, the whole picture of disaster management will become clear. This step we will mainly put our focus on the integration of the ontologies for processes and the data. The result will be a complete ontology for the domain of disaster management.

4. Implementation and test of ontologies and modifications of ontologies.

- Implementation tools and architecture selection. At this stage, comparisons will first be made on different ontology representation languages, such as, DAMO+OIL, RDF, OWL (includes OWL Full, OWL Lite, OWL DL and OWL-S) and WSDL, to choose the best one (or several) to represent the ontologies in machine-processable language. Second, different existing environment will be compared, such as Protégé, WSMO studio [9], to implement the ontologies. After that, we will examine the existing knowledge reasoners, for instance MORE (Multi-version Ontology REasoner), RacerPro, WSDL reasoner and so on, to see if they are suitable for the use of the PhD project.
- Implementation test of ontologies for ED (existing data) integration. After having the vocabulary (ontologies) for existing data, we will implement the ontologies for ED and carry out a test to see if the vocabulary (ontologies) works well for the integration of the existing data. Modification will be made according to the result of the test.
- Implementation and test of ontologies for ExOp data (existing and operational data). We will implement the ontologies for ExOp and carry out a test to see if the vocabulary (ontologies) works well for the integration of existing data and operational data. Modification will be made on the vocabulary (ontologies) for existing data and operational data according to the result of the test.
- Implementation and test of domain ontologies prototyping. The result of the previous step will be a series of queries. The queries will be used to test the domain ontologies application. An evaluation of the domain ontologies application will be made according to the evaluating criteria. Modification of domain ontologies. According to the result of the test, a modification will be made on the domain ontologies.

5. Evaluation of this project. Compare the results from the demos and prototypes that have been built using the ontologies with the current solutions that are used without ontologies. Find out the added value of formal semantics. (This process will carried out together with the previous process — implementation and test of ontologies and modifications of ontologies).

6. Final demo building.

    - Scenario and criteria selection. This step selects a scenario from disaster management. A set of queries will be selected from disaster management to show the use of the system. And a set of evaluation criteria, such as the robustness, the time-consumption and etc., will be chosen to evaluate the use of the domain ontologies. The result will be a set of queries and evaluation criteria for test the domain ontologies application. This scenario will be use primarily for GDi4DM and HUMBOLDT (perhaps others).
    - Prototyping of domain ontologies. A web services that uses ontologies will be made available to provide different services, such as format conversion, schema mapping, ontologies access, thesaurus access, knowledge service, query mediation and so on. A prototype application combining and using the service will be carried out between the users and the data to show the use of the domain ontologies. The result is an application (system) using the domain ontologies to facilitate the discovery of data. The scenario and criteria selection will be used to demonstrate the use of the system.

7. Final preparation.

    - Final release of domain ontologies. At this step, a final work will be made on the release of the domain ontologies for disaster management. And evaluation of the added values of formal semantics will be made.

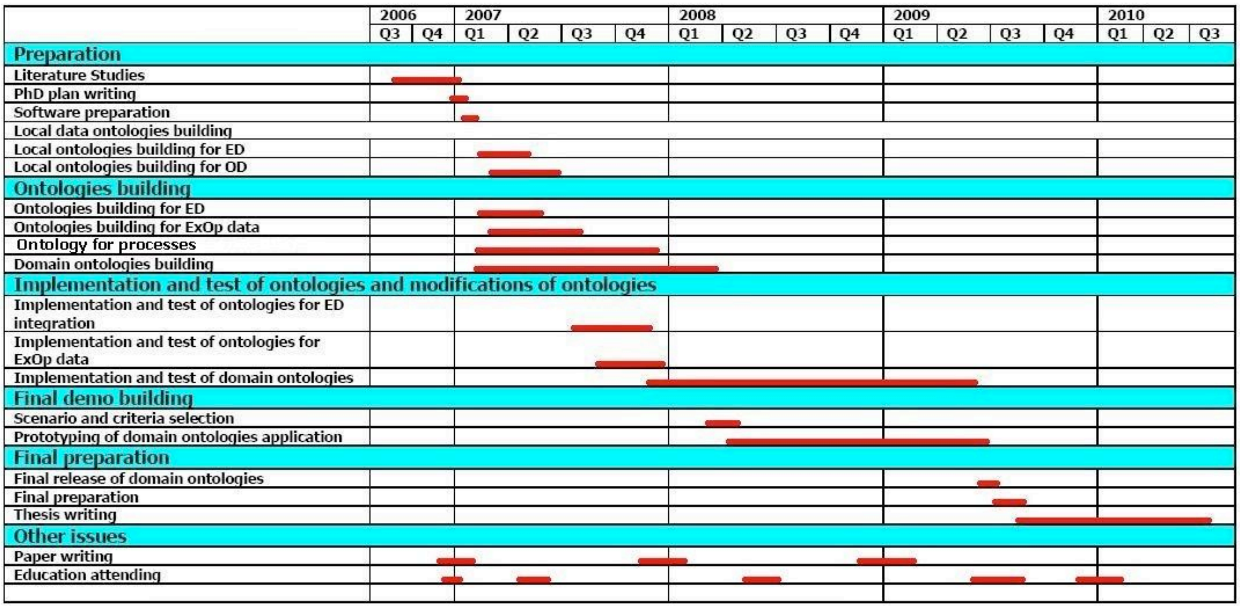


Figure 4.1: Phd Plan

- Final preparation. At this stage, the PhD research is finished. And this is the last stage for the preparation of the writing of PhD thesis.
- Thesis writing. The last period of PhD research will be mainly spent on the writing of the PhD thesis. And after the submission of the PhD thesis, one or two months will be spent on writing other articles (this depends on the time and other factors).

8. Other issues.

   Writing of paper and chapters of PhD thesis will be carried out as the research progresses. And attendance of education and conference will be carried out during the research period.

The above is a list of research agenda. And the time for writing paper and reports, visits to related institutes, universities and organisations, and attending education and conferences is also included (but not specified). All these mentioned activities depend on the relevance and the time related to the research.

4.2 Organisation Issues

1. Planned meeting with supervisors. Three scientific staffs are involved in the supervision work of the PhD research — a daily supervisor Sisi Zlatanova, a technical supervisor Marian de Vries and the promotor Peter van Oosterom. As it is important and helpful to track the work of the PhD research, a planned meeting agenda with supervisors is proposed and notes at meeting will also be made.

   - Once (or twice) per week meeting with Sisi Zlatanova to discuss the detailed issues on disaster management.
   - Once per two weeks meeting with Marian de Vries to discuss the technical issues of building ontologies and semantic technologies (esp. in the first year).

- Once per month meeting with all the supervisors to report the process of PhD research.

2. Planned conferences. In order to get to know the start-of-the-art advancements and show the achievements of the PhD research, an agenda to attend several conferences is also planned, which includes (this is not a exhaustive list of all planned conferences):

   - Gi4DM (website: `http://commission4.1uphost.net/gi4dm.html`), to use the latest state-of-the-art space-based geomatics technologies to understand the dynamic earth processes and geo-hazards (in the year of 2007 and 2008).
   - ISCRAM (website: `http://www.iscram.org/`), an international community on information systems for crisis response and management (in the year of 2008 or 2009).
   - SebGIS (website: `http://www.cs.rmit.edu.au/fedconf/`), Semantic-based Geographical Information Systems, aims at "discussing views on how to integrate semantics into current geographic information systems, and how this will benefit the end users".
   - AGILE (website: `http://www.agile-online.org/`), "promotes academic teaching and research on Geographic Information Science by representing the interests of those involved in GI-teaching and research at the national and the European level, and the continuation and extension of existing networking activities" (in the year of 2007 or 2008).
   - SDH (website: `http://www.igugis.org/conference.htm`), Spatial Data Handling. The International Symposium on Spatial Data Handling (SDH) is the premier long-running forum in geographical information science, providing a prestigious outlet to geographers, cartographers, computer scientists and others in this rapidly developing multidisciplinary area.
   - ISWC (website: `http://iswc.semanticweb.org/`). The International Semantic Web Conference is a major international forum at which research on all aspects of the Semantic Web is presented.
   - TIEMS (website: `http://www.tiems.org/index.php`), an non-profit society informs and educates the public in all areas of emergency management, publishes periodicals, sponsors networking events (in the year of 2008 or 2009).
   - ISPRS (website: `http://www.isprs.org/`), International Society for Photogrammetry and Remote Sensing. It is "an international NGO devoted to the development of the international cooperation for the advancement of knowledge research, development and education in the Photogrammetry, Remote Sensing and Spatial Information Sciences, to contribute to the well being of humanity and sustainability environment" (in the year of 2008 or 2009).
   - UDMS, Urban Data Management Society (website: `http://www.udms.net/`), organizes "international symposia at various locations in Europe in order to promote the development of information systems in local government" (in the year of 2008 or 2009).
   - COSIT, Conference On Spatial Information Theory

   All the planned conferences are not limited to the above list, some other conferences related with geo-information, disaster management and semantic web will also be attended considering the relevance and time w.r.t. the PhD research.

3. Planned educations. An education agenda for PhD research is also made, which includes (but not limited to):

   - Attendance of PhD schools, for instance VESPUCCI (summer institute on Geographic Information Science).

- Lectures on ontologies and semantic web (Ontology engineering) in Free University of Amsterdam and lectures on GIS (such as, Geo Database Management Systems, Geo Information Infrastructure Technology and so on) in Delft University of Technology.
- Some on-line courses are also included, for instance, Semantics and Ontologies in Geographic Information Services from the Technical University of Vienna.

4. Planned visits to other universities, institutes, or organisations. In order to get cut-the-edge technologies and research results, a planned visiting is also considered, which includes (but not limited to):

   - Universities and other research institutes, for instance the Free University of Amsterdam, which specializes in ontologies, ITC (International Institute for Geo-information Science and Earth Observation), whose knowledge field is geo-information science, earth observation and natural hazards.
   - Organisations. Since the PhD research needs much expertise in some certain areas, visits to some specialized organisations will help and advance the research, for instance, the fire brigade department, the police department, some mapping agencies and so on.
   - Companies. Some world leading companies are also busy the development of Semantic (Ontologies) technologies, such as Oracle. The GDI4DM project has cooperations with Geodan (website: `http://www.geodan.nl/`) and Nieuwland (website: `http://www.nieuwland.nl/`). Visits to these companies will also help the advancement of PhD research.

5. Planned Publications. A planned publications agenda is listed as follows (but not limited to):

   - Reports. This includes reports on the projects (GDI4DM and Humboldt), the progress, the development explanation and project results.
   - Papers. This includes paper submitted to workshops, symposiums and conferences. Each year two or three paper will be submitted to those mentioned events as specified above.
   - Journals. One or two journals in total in the four years of the PhD research.

References

[1] Agarwal, P. Ontological considerations in giscience. *International Journal of Geographical Information Science 19* (2005), 501–536.

[2] Arens, Y., Hsu, C. N., and Knoblock, A. Query processing in the SIMS information mediator. In *Readings in Agents*, M. N. Huhns and M. P. Singh, Eds. Morgan Kaufmann, San Francisco, CA, USA, 1997, pp. 82–90.

[3] Audi, R. *The Cambridge Dictionary of Philosophy.* Cambridge University Press, 1999.

[4] Bernard, L. Experiences from an implementation testbed to set up a national sdi. In *Proceedings of 5th AGILE conference on geographic information science* (2002), pp. 315–321.

[5] Borst, W., Akkermans, J. M., and TOP, J. L. Engineering ontologies. *International Journal of HumanComputer Studies 46* (1997), 365–406.

[6] Cox, S. Ogc: Observations and measurements. Tech. rep., Open Geospatial Consortium, 2006.

[7] Cutter, S. L., Richardson, D. B., and J., W. T. *Geographical Dimensions of Terrorism.* Routledge, 2003.

[8] DHS, U. Department of homeland security geospatial data model. http://www.fgdc.gov/fgdc-news/geo-data-model/ last accessed in June 2006, 2007.

[9] Dimitrov, M., Simov, A., Momtchev, V., and Konstantinov, M. Web service modeling ontology. Tech. rep., Ontotext Lab. / Sirma Group, 2007.

[10] Douglas, D. N. Developing spatial data infrastructures. Tech. rep., Global Spatial Data Infrastructure Association, 2000.

[11] DSTC, Gentleware, IBM, and Software, S. Omg's ontology definition metamodel. Tech. rep., IBM Sandpiper Software, Inc., 2005.

[12] Gruber, T. R. A translation approach to portable ontology specification. *Knowledge Acquisition 5* (1993), 199–200.

[13] Gruber, T. R. Toward principles for the design of ontologies used for knowledge sharing. *International Journal of Human-Computer Studies 43* (1995), 907–928.

[14] Guarino, N., Carrara, M., and Giaretta, P. An ontology of meta-level categories. In *KR'94: Principles of Knowledge Representation and Reasoning*, J. Doyle, E. Sandewall, and P. Torasso, Eds. Morgan Kaufmann, San Francisco, California, 1994, pp. 270–280.

[15] Hess, C., and de Vries, M. From models to data: A prototype query translator for the cadastral domain. *Computers, Environment and Urban Systems 30* (2006), 529–542.

[16] Jernst. What are the differences between a vocabulary, a taxonomy, a thesaurus, an ontology and a meta-model. Tech. rep., www.metamodel.com, 2003.

[17] Jones, E., and Botterell, A. Common alerting protocol. Tech. rep., OASIS Standard, 2005.

[18] Kassel, G. Integration of the dolce top-level ontology into the ontospec methodology. Tech. rep., Knowledge Engineering Team, University of Picardic Jules Verne, 2002.

[19] Klien, E., and Lutz, M. The role of spatial relations in automating the semantic annotation of geodata. In *COSIT* (2005), pp. 133–148.

[20] Lemmens, R. L. G. *Semantic Interoperability of Distributed Geo-services.* PhD thesis, Delft University of Technology, 2006.

[21] Lieberman, J. Geospatial incubator. Tech. rep., W3C Geospatial Incubator Group, 2006.

[22] Lutz, M., and Klien, E. Ontology-based retrieval of geographic information. *Internatioal Journal of Geographical Information Science 20* (2006), 233–260.

[23] M., G. F. Recognition of gis critical. *GIM International 21* (2007), 49.

[24] Manola, F., and Miller, E. Rdf primer. Tech. rep., World Wide Web Consortium, 2004.

[25] Masolo, C., Borgo, S., Gangemi, A., Guarino, N., and Oltramari, A. Ontolgoy library. Tech. rep., Wonder Web, 2003.

[26] McGuinness, D. L., and van Harmelen, F. Owl web ontology language overview. Tech. rep., world wide web consortium, 2004.

[27] McQueen, B., and McQueen, J. Iso tc204 (document n271). Tech. rep., ISO, 1999.

[28] N., G. Formal ontology, conceptual analysis and knowledge representation. *International Journal of Human-Computer Studies 43* (1995), 625–640.

[29] Neches, R., Fikes, R., Finin, T., Gruber, T., Senator, T., and Swartout, W. Enabling technology for knowledge sharing. *AI Magazine 12* (1991), 36–56.

[30] Niles, I., and Pease, A. Towards a standard upper ontology. *Proceedings of the international conference on Formal Ontology in Information Systems 2001* (2001), 2–9.

[31] Parliament, E. Establishing an infrastructure for spatial information in the european community. Tech. rep., European Parliament and the Council, 2007.

[32] Probst, F. An ontological analysis of observations and measurements. In *Proceedings of 4th International Conference, GIScience 2006* (2006), M. Raubal, H. J. Miller, A. U. Frank, and M. F. Goodchild, Eds., vol. 4197, Springer, Berlin., pp. 304–320.

[33] Riecken, J., Bernard, L., Portele, C., and Remke, A. North-rhine westphalia: Building a regional sdi in a cross-border environment/ad-hoc integration of sdis: Lessons learnt. In *Proceedings of Proceedings of 9th EC-GI & GIS Workshop ESDI* (2003), pp. 25–27.

[34] Robertson, D., Giunchiglia, R., van Harmelen, F., Marchese, M., Sabou, M., Schorlemmer, M., Shadbolt, N., Siebes, R., Sierra, C., Walton, C., Dasmahapatra, S., Dupplaw, D., Lewis, P., Yatskevich, M., Kotoulas, S., de Pinninck, A. P., and Loizou, A. Open knowledge semantic webs through peer-to-peer interaction. Tech. rep., Open Knowledge, 2006.

[35] Sheth, A. Changing focus on interoperability in information systems: From system, syntax, structure to semantics. In *Interoperating Geographic Information Systems* (1998), M. Goodchild, M. Egenhofer, R. Fegeas, and C. Kottman, Eds., pp. 5–30.

[36] Smith, B. The basic tools of formal ontology. Tech. rep., Institute for Formal Ontology and Medical Information Science, Saarland University, 1998.

[37] Smith, B. Basic formal ontology and medical ontology. Tech. rep., Department of Philosophy, State Univerisity of New York at Buffalo, 2003.

[38] Smith, B., Kumar, A., and Bittner, T. Basic formal ontology for bioinformatics. Tech. rep., Alexander von Humboldt Researcher, University of Leipzig, 2005.

[39] Snoeren, G. User requirement for a spatial data infrastructure for emergency management. Master's thesis, GIMA, 2006.

[40] Stubkjær, E. Integrating ontologies: Assessing the use of the cyc ontology for cadastral applications. In *ScanGIS* (2001).

[41] Studer, R., Benjamins, R., and Fensel, D. Knowledge engineering: principles and methods. *Data and Knowledge Engineering 25* (1998), 161–197.

[42] Swartout, B., Patil, R., Knight, K., and Russ, T. Toward distributed use of large-scale ontologies. In *Proceedings of the 10th Knowledge Acquisition for Knowledge-Based Systems Workshop* (Banff, Canada, 1996).

[43] Uitermark, H. *Ontology-Based Geographic Data Set Integration.* PhD thesis, University of Twente, 2001.

[44] Uschold, M., and Gruninger, M. Ontologies: principles, methods and applications. *Knowledge Engineering Review 11* (1996), 93–155.

[45] van Binnenlandse Zaken en Koninkrij, M. Grip levelksrelaties. http://www.handboekrampenbestrijding.nl/ last accessed in July 2006, 2006.

[46] van Borkulo, E., Scholten, H. J., Zlatanova, S., and van den Brink, A. Decision making in response and relief phases. In *Geo-information for Disaster Management,* (2005), pp. 47–53.

[47] van Oosterom, P., Lemmen, C., Ingvarsson, T., van de Molen, P., Ploeger, H., Quak, W., Stoter, J., and Zevenbergen, J. The core cadastral domain model. *Computers, Environment and Urban Systems 30* (2006), 627–660.

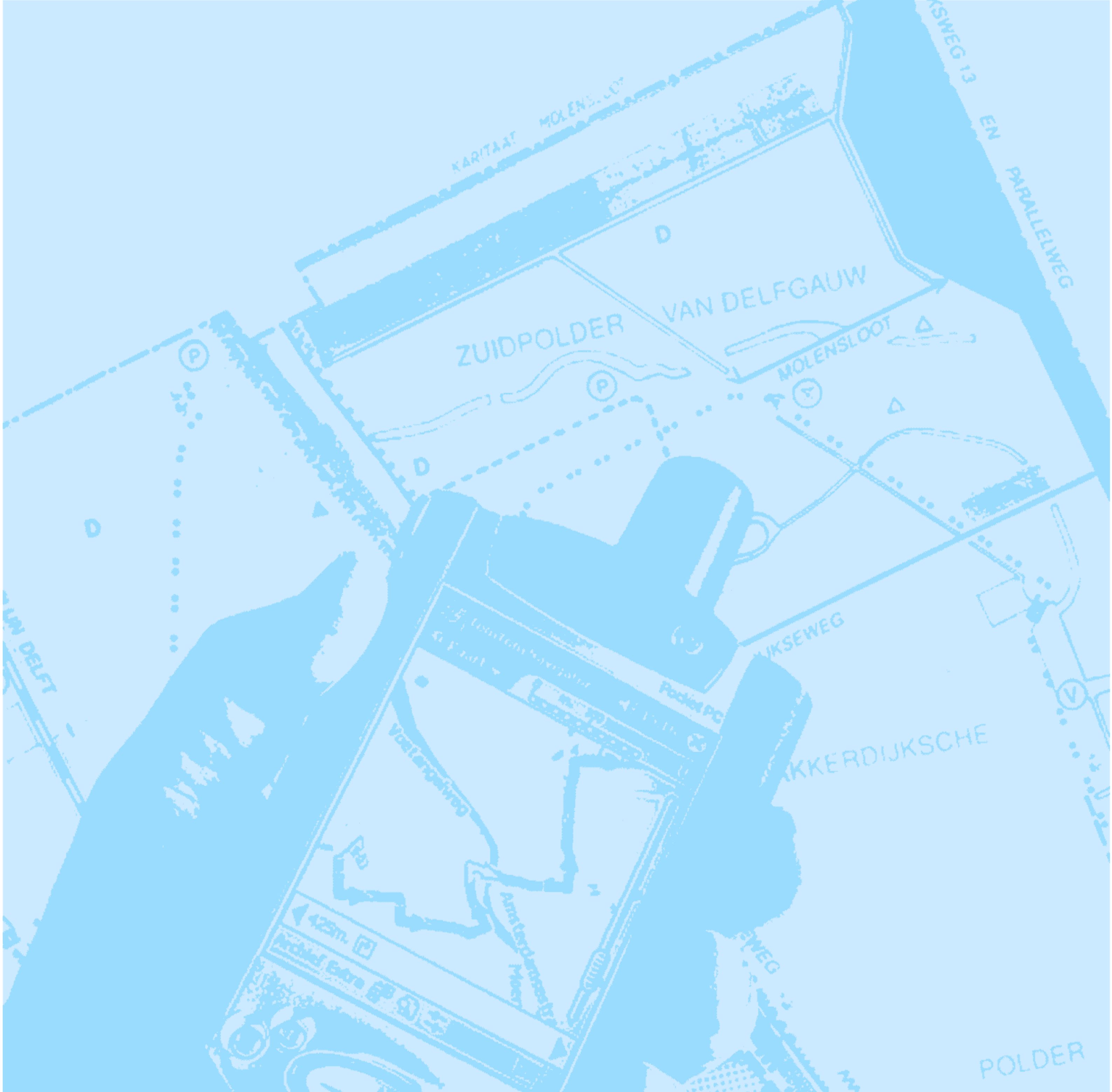
[48] Visser, U., Stuckenshmidt, H., and Schlieder, C. Interoperability in gis-enabling technologies. In *Proceedings of AGILE Conference on Geographic Information Science* (2002), pp. 25–27.

[49] Wache, H., Vgele, T., Visser, U., Stuckenschmidt, H., Schuster, G., Neumann, H., and Hbner, S. Ontology-based integration of informationa survey of existing approaches. In Stuckenschmidt, H., ed., IJCAI-01 Workshop: Ontologies and Information Sharing, 2001.

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.
45. GISt Report No. 45, W. Tegtmeier, 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, Delft, 2007, 40 p.


ZUIDPOLDER
VAN DELFGAUW
MOLENSLOOT
PARALLELWEG
AKKERDIJKSCHE
POLDER