SEMANTIC DATA INTEGRATION IN A MULTIPLE REPRESENTATION ENVIRONMENT

J.E. Stoter[a] , R.L.G. Lemmens[a], B. Kőbben[a] and N. J. Bakker[b]

[a] ITC, International Institute for geo-information science and earth observation, P.O. Box 6, 7500 AA Enschede, the Netherlands, stoter@itc.nl

[b] Topografische Dienst Kadaster, Bendienplein 5, 7815 SM Emmen, The Netherlands, nbakker@TD Kadaster.nl

Working Groups II/3 and II/6

KEY WORDS: Semantic data integration, Multi Representation Database, Generalisation, ontologies, machine ontology

ABSTRACT:

Traditionally National Mapping Agencies (NMAs) produce maps at different scales. Since automatic generalisation is not yet possible (it will at least take years) NMAs have to maintain their data sets in a multiple representation environment. Important principles of this environment are 1) the data sets at the different scales are semantically integrated, 2) objects in the different scales representing the same real world objects are linked and 3) updates in the largest scale data set are automatically propagated to the smaller scales. This paper describes the considerations for semantic data integration of the topographic database of the Dutch NMA (TD Kadaster). In collaboration with ITC a generic data model is being designed in which the data models of the different scales are integrated. This model supports topographic data sets at a scale range from 1:10k to 1:500k. The results of this data model integration that were obtained so far are reported in this paper. The difficulties that are met are related to using existing data sets in the data integration process. For example the classifications of the different data sets (i.e. map legends) are not synchronised, since this was not needed when producing paper maps. Also the generalisation rules that are verbally described in cookbooks to support the interactive work of cartographers seem not to be in accordance with the classification schemas of the existing data sets. Apart for internal processes, semantic data integration is also required for the exchange of geo information with other geo information suppliers. This paper describes how semantic data integration of Dutch topographic data with the outside national world is assured by embedding the topographic data model in the Dutch standardised data model for geo information. The embedding is modelled using Unified Modelling Language. The model is limited in some aspects since GML is used as exchange standard. Therefore this paper considers the possibilities of a new technology in the semantic data integration process of different data sets: machine ontology using the Web Ontology Language (OWL).

1. Introduction

There are two types of standards: standards for technology and standards for semantics. Standards for technology support the communication and interaction between different software systems by means of common interfaces. Standards for semantics support the unambiguous understanding and use of information which are collected by different organisations, at different time periods and for different purposes. This makes the semantic web possible which is a web of information that machines can process without the interactivity of human (Daconta et al, 2003). Originally the technological standards, such as the ones set by OGC and ISO, were used for the interoperability of systems and standards for semantics were used to make people behind the systems understand the information that is being presented. However in the future major parts of information will be used by systems before it is presented to humans, e.g. in generalisation. Humans can understand concepts by using the contexts of the concepts (what is the domain, who is the information supplier). However for systems it is necessary to make all knowledge explicit. This can be done by formalising the semantics, i.e. by defining objects and their relationships.

This paper describes the semantic aspects that need to be considered to support a full automatic generalisation process (or at least as automatic as possible) within a National Mapping Agency (NMA). It reports about the semantic data integration which is necessary to convert the traditional production lines of topographic mapping into a production line in which databases at different scales are maintained in a multi representation environment, including the links between multi-represented objects and the support of update propagation. The semantic data integration is the aim of our research project, which is a cooperation of the Dutch NMA Topografische Dienst Kadaster (TD Kadaster), ITC, Wageningen University and Research Centre and Delft Technical University. In our project we are defining an overall data model covering all scale ranges from 1:10k to 1:500k. This report describes the tasks of ITC in the project: designing the integrated data model. At this stage we are only considering the semantic data integration. Geometric data integration and cartographic generalisation will be the next steps in making the generalisation process fully automated (or as automated as possible).

In section 2 the current production line of TD Kadaster is described to show its ineffectiveness. Section 3 starts with an overview of integration problems within different NMAs, then it describes the improvement of the production line of TD Kadaster based on 1) the definition of a base data model and 2) the semantic data integration of all scales. Semantic interoperability is not only important for the internal working processes but also to be able to exchange data sets with others. The practice of TD Kadaster is used to show how semantic interoperability with other (Dutch) geo information sets can be assured (section 4). Section 5 finally describes a new technology for the semantic integration of different data sets: the Web Ontology Language (OWL), which can support the business of NMAs. As will be seen from this section the

machine ontology offers new possibilities for the semantic web. This paper ends with conclusions.

2. Current production line of TD Kadaster

The TD Kadaster is responsible for producing topographic maps and databases at different scales (1:10k to 1:500k). For these products they maintain vector data sets at different scales. By using automatic symbolisation the cartographic products can be produced. Currently the vector data sets are maintained separate from each other while updates at the different scales are performed independently and, for the main part, interactively. For example the 1:50k is updated by projecting a 1:10k data set on the screen and the updates are performed by hand using some automatic tools. The smaller scales (1:100k, 1:250k and 1:500k) are even updated independently from the larger scale data sets, since they have a different update cycle. Of course the most recent information from the larger scale data sets is used, but because the map sheets and the update cycles do not always coincide, the information from the larger scale data sets that is used is a mix of recent and old information. Occasionally the old information is completed with new information from the field which is specifically collected for the smaller scales. Obviously this is not an optimal situation: the production is expensive and it takes considerable time to have updated versions of all scales; updates are not available at the same time in the different data sets which leads to inconsistencies. In addition more flexible data sets would provide the possibility to produce demand driven products instead of the supply driven, predefined topographic data sets. Therefore the TD Kadaster is currently investigating the possibilities to improve their production line.

3. Improving production line by integration of datasets

The most optimal solution would of course be only maintaining one base data set from which any selection of objects at any scale could be dynamically generalised. A less advanced solution would still only maintain one base data set while generalising whole data sets at the predefined smaller scales. The advantage would be not to have to maintain data sets at different scales. Since full automatic generalisation processes do not exist, the TD Kadaster had to find a more feasible solution, i.e. maintain all vector data sets as object oriented databases in a multi representation environment and implement (as) automatic (as possible) update propagation using automatic generalisation.
Considering improvement of the production line it is also important to look at the specifically Dutch developments of authentic base registrations which are defined by the Dutch government. The aim of the base registrations is to collect information only once and use it many times. This will be enforced by law during the coming years. Governmental authorities will be obliged to use the base registrations that are defined as such while citizens and companies can be requested only once to submit their information. In the proposed new law the base data set of the TD Kadaster (1:10k) has been assigned as authentic registration for topography. This lays down specific requirements for the quality and update cycle of the data set, enforcing the TD Kadaster to look at a more optimal production line. Especially since it is expected that the smaller scales will soon follow as authentic base registrations once the base registrations that have currently been defined become practice. One of the conditions to support authentic base registrations at all scales will be that the automatic propagation of updates is realised through the whole scale range. This requires integration of the data sets at the different scales.

3.1 Integration of data sets at different scales in NMAs

Also other NMAs meet this problem see for example (Féchir and de Waele, 2005) in which it is described how to integrate existing datasets at scales 1:10000 and 1:50000 in order to support future automatic update propagation in a Multiple Representation DBMS. In (Stoter, 2005a and Stoter, 2005b) the trends in a number of NMAs is highlighted. All NMAs meet the problems of applying new geo-information technologies to existing datasets. To avoid integration problems of existing data sets in Germany the 1:50.000 is derived once from the 1:10.000. In some Bundesländer the links between the different scales are maintained. France is also working on establishing links between objects in different data sets and have achieved results in the midscale. Sweden is redesigning its data models for all scales while trying to keep the coherence between. An extra problem that Ireland is faced with is that the scale and semantics in one datasets (the base dataset) varies depending on the type of area: 1:1k in urban area, 1:2.5k in non-urban area and 1:5k in remote areas. At this moment the different specifications are being reviewed and compared in order to indicate differences and similarities. For example in the current dataset it occurs that the 1:1k data contains less information for some themes than the 1:2.5k and 1:5k data.

3.1 Semantic integration within TD Kadaster

Remodelling the base scale

The first step in improving the production line was to remodel the largest scale data set to fulfil today's technical and user´s requirementsneeds. The new version of the 1:10k data set is called TOP10NL and the model (version 2.3) was established in February 2005. An extensive description of the TOP10NL model can be found in Bakker (2005). Here we give a short summary. Designing the data-model and its structure the following user´s requirements were taken into account:

- Object-oriented data model with unique IDs,
- Structured attributes with the possibility to add new attributes,
- Based on international standards (ISO, OGC),
- Change-only data instead of complete update delivery,
- Availability of object life cycle information,
- Metadata on object level,
- Distinction between Digital Landscape Model and Digital Cartographic Model,
- Seamless database,
- Integrating the former TOP10vector (roads as polygon) and TOP10roads (roads as centre-line),
- Prepared for automatic generalisation and scale-less database,
- Linkage with user data models (e.g. environment, water) and user data sets.

In the new model (figure 1) a collection of topographic base objects has been defined.

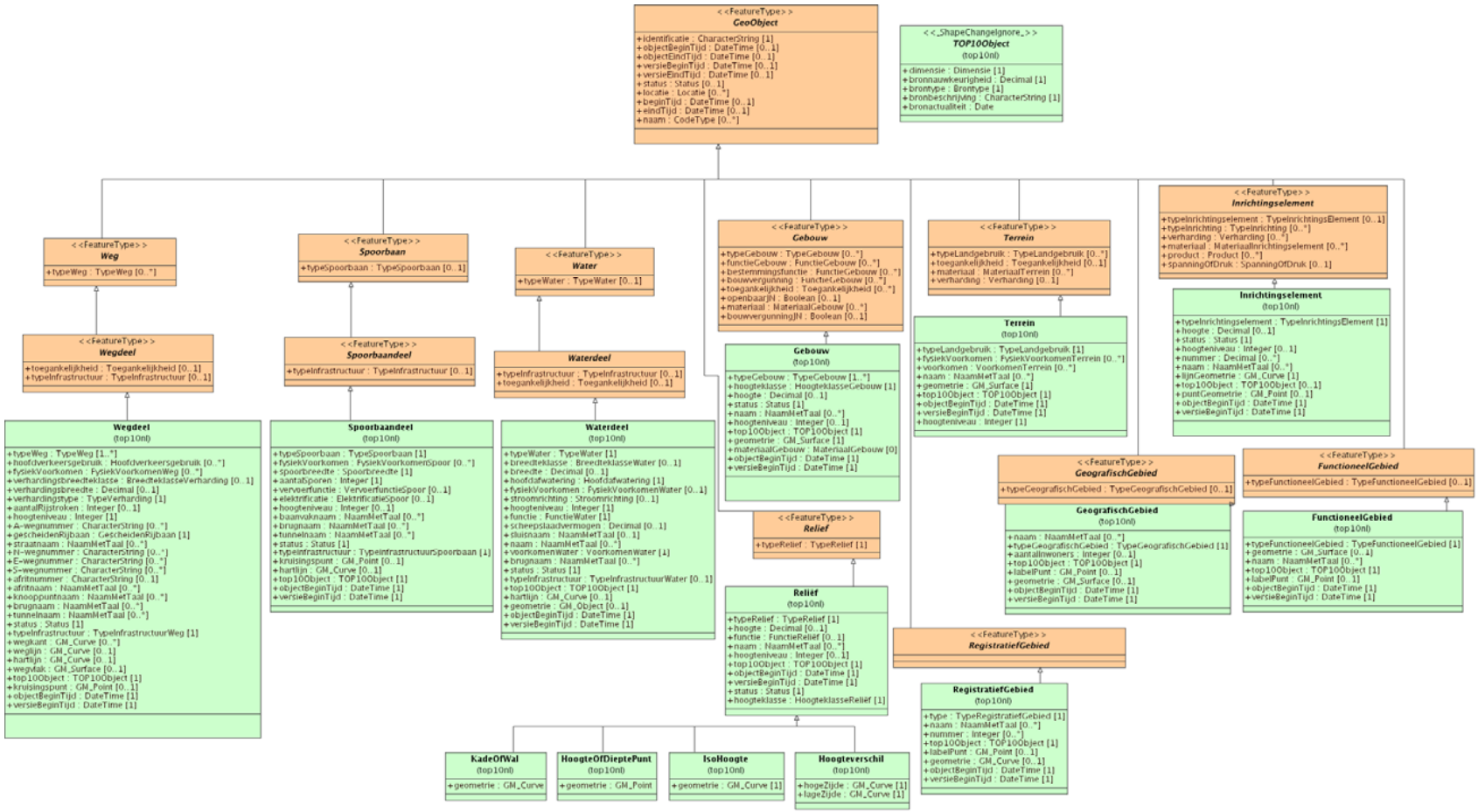

Figure 1: UML model of TOP10NL

The features that were already stored in the current database (TOP10vector) were rearranged according the new defined object types:

- Road segment
- Railroad segment
- Water segment
- Terrain
- Building
- Specific terrain element/ construction
- Administrative area
- Geographic area
- Functional area

The data model is modelled in the Unified Modelling Language (UML). A detail of the data model in UML is shown in figure 1. The topographic base objects have unique identifiers, meta characteristics and describing attributes. They also have temporal attributes in order to monitor history in the database. Geometries of objects are points, straight lines or polygons consisting of boundaries of straight lines. There are objects which contain more than one geometry, e.g. a road can be defined with a polygon, its centre lines and the sides of the roads. Complex objects exist, for example a collection of road segments which together form a specific road (E 32). These complex objects are not explicitly modelled but can be found by looking for common characteristics (e.g. road name). Complex objects can also be used to represent a collection of objects as one object at a higher level of detail, e.g. 'wood' in a mid-scale database can be a collection of three types of land use in the TOP10NL: coniferous wood, deciduous wood, and mixed wood.

At the moment all 1:10k data sets are converted into the new TOP10NL model, which is expected to be ready at the beginning of 2006. An example of TOP10NL data is shown in figure 2.

Figure 2 Fragment of the TOP10NL dataset

It is interesting to see how data models which look ideal in "laboratory" conditions are not easy to apply to existing data sets. The digital data were originally collected to support map production, now they have to be improved and enriched to be able to fulfil the new (object-oriented) requirements (consistent crossing of road and water segments since the new data model supports 3D features; correct classifications; perfect geometries to support the topological model, no spaghetti or dangling arcs etc.).

A generic multi scale data model

The next step towards automatic generalisation of updates is to redesign the data models of the smaller scales into object-oriented data models. To support the whole generalisation cycle, a generic multi-scale data model for the scale range from 1:10k to 1:500k (and smaller) is being designed. This model will, as the TOP10NL model, be based on the Basismodel Geoinformatie (see section 4), OGC and ISO specifications. From this data model the smaller scale data models can be derived using model generalisation.

Results until now

To come to a generic model, the map legends of TOP10vector, TOP50vector, TOP100vector, TOP250vector and TOP500vector need to be compared with the TOP10NL classes in order to identify objects in the smaller scales which are not included in the TOP10NL model. The currently available data sets at the different scales were evaluated and gaps between these scales and the TOP10NL were indicated. However a more optimal solution would be to reconsider the data models at the smaller scales as was done with the TOP10vector before indicating the gaps between TOP10NL and the smaller scales. This is underlined by the fact that at first glance the different scales do not seem to be synchronised: they are not in accordance (see also Knippers and Kőbben, 2003). The urge to synchronise the classifications of the different scales has become larger now databases are fundamental for the production line whereas in former times, there was no need at all to link the different legends. Revision of the data models at the smaller scales will require a tight cooperation between practice (cartographers who make the maps) and the scientists (who probably have more ambitious ideas that are only tested in laboratory situations, not limited by current practice and that will work on the long term). It will definitely be a challenge to get the scientists and the cartographers in agreement. It should be noted that our project will give suggestions and recommendations for the semantic improvement for the data models at the different scales in order to have them semantically integrated.

The next step is adjusting the TOP10NL model in order to include classes that occur at the smaller scales but do not occur in TOP10NL. This will yield the TOPNL data model.

Semantic data integration is based on generalisation methods such as the *selection* of object classes and attributes, the *reclassification* of objects, and the *typification* of objects as point, line or polygon. Based on the TOP10NL data model the definition of a selection schema for every small scale data set can be made. The schema should indicate whether an object at the smaller scale needs to 1) remain unchanged, 2) be omitted, 3) be reclassified and how it needs to be classified (e.g. "block of flats" in the TOP10NL data model is reclassified into a class area with "high objects" in the TOP250vector) 4) be typified. The classification schemas form the conceptual part of the generalisation process. Based on the selection schemas the databases at the smaller scales can be derived from TOP10NL.
A start was made of making the selection schemas. However since the different scales are at this moment not synchronised, the selection schemas are sometimes inconsistent and incoherent. This once again emphasis the fact, that well considered data models are indispensable, also at the smaller scales. It is therefore probably better to design the selection schemas and the new data models at the same time (i.e. base the

classifications at the smaller scales on the classification in the large scale data set).

The final step in the modelling phase is to set up generalisation rules for every object class. These generalisation rules are applied after symbolising the reclassified objects and are realised through operations such as *simplifying, enlarging, displacement, aggregation, deletion.*
The TD Kadaster has written generalisation 'cookbooks' to generalise the data sets at the smaller scales. The cartographer can use the rules in this cookbook to produce a small scale data set. For automatic generalisation these rules need to be formalised, which will not be an easy task, since a computer cannot make decisions based on the context of objects unless the context is formalised. Which means every exception (which can be many in generalisation) needs to be encoded into computer rules. General generalisation rules will be defined at class level. Specific generalisation rules will be defined at object level and it should be possible to indicate if a general rule, defined at class level, leads to a conflict for a specific object.
For example the priority rules for displacement defined at class level can be defined in the following order:

- Water segment
- Railroad segment
- Road segment
- Terrain
- Building
- Specific terrain element/construction

A general rule for displacement is that an object high on a higher class level will not be replaced because of an object on a lower class level.
When studying the cookbooks we came to the conclusion that the TOP10vector classifications schemas and the classifications that are used in the generalisation cookbook are not harmonised. So also in this phase one should not apply new technologies to do the same thing. The generalisation cookbooks have to be reconsidered at the fundamental level, based on the new data models and based on the new selection schemas.

4. Data integration with the outside (national) world

During the final stage of finishing the data model for TOP10NL a more general data model for geo information in the Netherlands was completed. This data model is the national standard for modelling geographical data and is a data model which provides the concepts, definitions, relations and general rules for exchanging information on objects which are related to the earth surface. The model is based on NEN3610 (originating from 1995, NEN is the Dutch Normalisation Institute). The old model was not in accordance with the international standards from ISO and OGC and therefore it was reconsidered in 2004/2005 under the name *Basismodel Geoinformatie* (Base model for Geo Information NEN3610 version 2/2005) (Aalders and Reuvers, 2004). The aim of this model is to have common definitions for geo objects (objects that occur in the terrain of which the location and geometry are fixed within a certain time period) focusing on the interoperability of information systems. Also objects which are not visible in the terrain are included such as cadastral parcels, administrative boundaries..
The new version of the model was designed by RAVI (Dutch organisation for geo information in the public domain) in collaboration with ICTU (Dutch organisation for information and communication technology in the public sector), Wageningen University and Research Centre, Technical University of Delft and the Netherlands' Kadaster. To achieve commitment many users were consulted. Within the model the geo objects are described with identifiers, descriptive, geometrical, temporal and meta data attributes. To be able to link to geo information outside TD Kadaster the TOP10NL data model has been embedded in the Basismodel Geoinformatie. This has been achieved by defining every class of TOP10NL as subclass of a NEN3610-class (see figure 1): every TOP10NL object is a NEN3610 GeoObject. By this all TOP10NL objects inherit the properties of the NEN 3610 GeoObject.

All TOP10NL objects have general characteristics in common, which are defined as attributes of the TOP10Object. It would be most optimal if every TOP10NL class would be a subclass of the TOP10NL object. However since GML has been chosen as exchange format and since in GML every class can have only one super class (multiple inheritance is not allowed) the Top10Object attribute has been added to all TOP10NL objects to assure that all TOP10NL objects share the TOP10NL characteristics such as dimension, source type, accuracy (see also Quak et al., 2005).
For many attributes in TOP10NL, a list with possible values is used. These coding lists are integrated with the NEN3610/2005 coding lists. The general coding lists of NEN3610 can be extended with TOP10NL specific codes. For example the TOP10NL model defines more types of forests than the NEN3610 data model.
In the future geo information providers and other sectors will make their own extensions of the NEN3610 model, by which geo information can be easily exchanged. In a sector model, (which is a model that covers the objects, relationships etc for the information in a specific application area) attributes of the NEN3610 model can be refined (e.g. in the NEN3610 the 'status' attribute is optional whereas it is obligatory in the TOP10NL model). However an obligatory attribute in the NEN3610 model can never be made optional in a sector model.

5. Machine ontologies for semantic integration in the spatial domain

In addition to the UML representation and the XML (or GML) schema, which could be considered as representations of ontologies, the semantics of data models can be captured in machine ontologies. These are formal, machine accessible representations of knowledge that are used for inferring intra- and inter-information resource relationships. Recent work within the Semantic Web community has resulted in W3C recommendations of two knowledge representation languages, i.e., the Resource Description Framework (RDF) and the Web Ontology Language (OWL) (McGuinness and van Harmelen, 2004). OWL draws upon the formal theory of Description Logics, which has roots in first-order predicate logic and provides highly expressive concept-forming constructs (Baader et al., 2003). These can be used to explicitly state the semantic relationships between concepts in or across data models in a more expressive way than the class hierarchies in current XML schema. The support of multiple inheritance in OWL based ontologies allows for a more flexible modelling of relationships then the 'straightforward' relationships in GML schemas. Therefore OWL enables to model more complex relationships between classes of different data models, which is specifically helpful when different data sets are different representations of reality but not of a different level of detail.
Formal ontology will push forward the geoICT community as argued in (Klien and Probst, 2005). They present an ontology

application example and derive from this a requirement specification for geospatial ontologies and the ontology architecture they are embedded in. They emphasis the importance of domain ontologies in the context of geospatial web service environments and they claim that the lack of a supportive environment for ontology engineering and maintaining decelerates the efficient use of ontologies in the GI community. Taking into account the requirements they identify a research action line which will help to establish such an environment.

Also NMAs can benefit from having their data models defined in ontologies since formal ontologies can overcome the problem of semantic differences between data sources. The electronical trading of information can only be successful if data and information from different sources can be shared and processed automatically as well as by people. Without ontologies different systems lack the ability to discern the meaning behind the shared data making true data integration impossible.

OWL and Ordnance Survey, UK

The Ordnance Survey in the United Kingdom also realised that real data integration creates opportunities to increase the use of Ordnance Survey's topographic information in ways that can go beyond the delivery of such data in the form of a map. They published several articles on research that they are doing in this area, see (Goodwin, 2005a; Goodwin, 2005b; Hart et al, 2004; Schwering and Hart, 2004; Greenwood and Hart, 2003; Hart and Greenwood, 2003). The goals of this research is long term but the intermediate benefits to Ordnance Survey are: better understanding and modelling of their data, improvements in their core database models, the development of intelligent web services, and understanding how data can be translated for many different user tasks. Furthermore, integrating Ordnance Survey data with other information sources may reduce both the time and cost of services, better enabling joined-up government and industry. In (Goodwin, 2005) the potential applications of a topographic ontology at Ordnance Survey are described, which are:

- sharing information to integrate and reuse knowledge and data across various applications
- making Ordnance Survey data more interoperable by translating between different semantics
- the 7000 rules that are currently used to check which combinations of geometry, form and function are valid, can be checked on consistency when expressed in OWL; a formally encoded rule about classification could aid the surveyor's decision making process; and when the rules need to be updated it is much more easier to modify the ontology and to check if the modifications are consistent.
- ontologies enable to integrate business, domain and policy knowledge in an intuitive information model.that can be easily tailored and adapted to user needs
- theintegration of semantic and spatial technologies support location based services. Queries can be made more easy and effectively when captured in protable and reusable from of an ontology than in using standard SQL

OWL and the Dutch TOP10NL data model

An example of a semantic relationship is the mapping between the classes of NEN3610 and TOP10NL for the concept `Hotel'. Due to the different way in which they model this concept, subclassing is not sufficient and a mapping of ontological properties is needed. Figure 3 portrays such a specific ontology mapping.

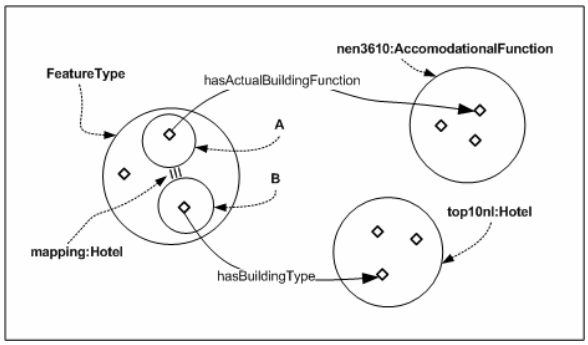

Figure 3. Schematic diagram of an example of ontology mapping between the NEN3610 and the TOP10NL data model.

Concepts are indicated by circles, instances by diamonds. An instance may have a relationship with another instance through a role. For example, a feature instance is related through the role `hasActualBuildingFunction' to `AccomodationFunction' (the *role filler*). Anonymous classes *A* and *B* are created to group instances with common role fillers. An equivalence mapping is created between *A* and *B* in Description Logic by a so called asserted condition as follows:

mapping:Hotel ≡ (exists *nen3610:hasActualBuildingFunction.nen3610:AccomodationalFunction*)
or (exists *top10nl:hasBuildingType.top10nl:Hotel*)

In the above statement the dot separates the role from the role filler. It has a corresponding OWL encoding that is used by a reasoner to infer the correct semantics of any instance (e.g., `Amstel Hotel') across both data models. It has to be noted that the above mapping example excludes potential other features with accomodational function, such as top10nl:Motel and top10:RecreationCentre. However, they can be easily added in the asserted condition. It is up to the information engineer, who integrates the models, to make the choices to include the right concepts in those mappings. More mappings and details of the reasoning process can be found in (Lemmens and De Vries, 2004) and (Lemmens, 2006). In addition, when a third model is involved (e.g. the data model of the domain of land use planning), and its concepts are related to either NEN3610 or TOP10NL, they are also connected to the other model through these mappings.

An integration of the data models can also take place in the realm of geometric representations. Feature types can be related to concepts in a geometry ontology. Geometry ontologies can be derived from the conceptual models such as ISO 19109 (ISO/TC211, 2003) and the OGC GML specification (OGC, 2003). The following Description Logic statement specifies the existence of point features within a data set:

exists (*hasFeatureType*.(exists *iso19109:hasSpatialAttributeType.gml:Point*))

Such statements can be used to compare different geometric representation concepts between geo data sets and services as the basis for geo information discovery, generalisation, integration, etc. Ongoing developments in ontology languages make it also possible to include rules that link an *antecedent* to

a *consequent*, where antecedent and consequent are OWL constructs. Ontology tools for editing, reasoning and software integration are subject to high pace developments and promise to bring ontology applications closer to implementation issues as described in this paper. In summary, machine ontologies can already be seen as a useful extension to XML schema when it comes to concept integration. Ontologies can also be used as data repositories, but for simple data carrier purposes XML documents and schema will suffice.

6. Conclusions

The Dutch TD Kadaster needs to establish a generic generalisation model to fulfil the customer's needs for scale less databases. Based on the digital topographic CAD files produced in the nineties of the last century the first attempts were made to develop an automatic generalisation process. However the database structure seemed not fitted for these processes. In the meanwhile a new object oriented database has been developed for the mid scale topographic data (1:10k): TOP10NL. In 2004 and 2005 the CAD database TOP10vector has been converted to the new object oriented data structure in TOP10NL. This new database is the base of a scale less data set.

The first step for the scale less data set is the development of a general data model based on TOP10NL which assures the semantic data integration of data sets at the different scales: TOPNL. This model will support the automatic generalisation process. The new generic data model will be developed for the whole scale range and used for defining the needed automatic generalisation processes.

The first steps to design the integrated model were reported in this paper.

From the first experiences of the semantic data integration process of different scales within the TD Kadaster it can be concluded that working with existing data sets that were originally developed for supporting map production yield problems. For example the classifications of the different data sets (i.e. map legends) are not synchronised, since this was not needed when producing paper maps. Also the generalisation rules that are verbally described in cookbooks to support the interactive work of cartographers seem not to be in accordance with the classification schemas of the existing data sets. So it is inevitable to reconsider existing data sets and data models, also at the smaller scales, in the semantic data integration process.

The old generalisation methods, which aimed at special (paper) products at different scales, should be reconsidered at a fundamental level, i.e. one should avoid making exactly the same product by applying new techniques. Apart from improving the data models to make the semantic integration possible, the data models should also be adjusted to new user needs.

Apart from the semantic data integration needed to support the internal processes of NMAs, semantic data integration is also required for the exchange of geo information with other geo information providers. This paper described therefore also how semantic data integration of Dutch topographic data with the outside national world is assured by embedding the topographic data model in the Dutch standardised data model for geo information: *Basismodel Geoinformatie*. The embedding is modelled using UML.

The model is limited in some aspects since GML is used as exchange standard. Therefore this paper considered the possibilities of new technology in the semantic data integration process of different data sets: machine ontology using the Web Ontology Language (OWL). As shown in the example this language offers good possibilities for modelling the relationships between objects from different data bases that do not have a pure hierarchical relationship, i.e. objects that are different in semantics but not necessarily in the level of detail.

In the geo information world a lot of effort has been focused on the interoperability of systems, i.e. developing technological standards. However with more players in the geo information market, (more users and more suppliers) the interoperability should now also focus on the exchange of information, i.e. computers should be able to understand the information within the context that it is meant and should be able to integrate information from different sources. This will yield important steps forward for the semantic web. Further research on the technology as shown in this paper is therefore needed.

References

Aalders, H.J.G.L and M. Reuvers, 2004, Het begin van een nieuw normentijdperk, in: Geo Info, Volume 2004, Number 1.

Aalders, H.J.G.L, J.D. Bulens, L.A.E. Vullings, P.A.L.M. Janssen, P.J.M. van Oosterom, M.Reuvers, C.W. Quak and H.T.J.A.Uitermark, 2004, Symposium Geo-Informatie Nederland 'De Nieuwste Ontwikkelingen in de Geo-informatie', Ede, June 11 2004, pp. 39-44

Baader, F., Calvanese, D., McGuinness, D., Nardi, D., and Patel-Schneider, P., Eds. The Description Logic Handbook; Theory, Implementation and Applications. Cambridge University Press, 2003.

Bakker, 2005, Developing a new geographical object database. Experiences from idea to delivering datasets. In: Proceedings of the 22nd International Cartographic Conference (ICC 2005), A Coruña, Spain, 9 - 16 July 2005

Daconta, M.C., L.J. Obrst and K.T. Smith, 2003, The Semantic Web: A guide to the future of XML, Web Services and Knowledge Management, ISBN:0-471-43257-1

Féchir, A. and J. de Waele, 2005, Institut Géographique National – Nationaal Geografisch, the future production of generalised maps at IGN Belgium, 8th ICA WORKSHOP on Generalisation and Multiple Representation, A Coruña, July 7-8th, 2005

Goodwin, J., 2005, Experiences of using OWL at the Ordnance Survey, presented at the workshop: OWL: experiences and directions, Galway, Ireland, November 11-12 2005.

Goodwin, J., 2005, What Have Ontologies Ever Done For Us – PotentialApplications at a National Mapping Agency, presented at the workshop: OWL: experiences and directions, Galway, Ireland, November 11-12 2005.

Greenwood, J. and Hart, G., 2003. Sharing Feature Based Geographic Information - A Data Model Perspective. Geocomputation.

Hart, G., H. Mizen and S. Temple. 2004. Tales of the River Bank: An overview of the First Stages in the Development of a Topographic Ontology. 12th Annual Conference on GIS Research UK (GISRUK) pp 179-186.

Hart, G. and J. Greenwood., 2003. A component based approach to geo-ontologies and geodata modelling to enable data sharing. Proceedings of the 6th AGILE, April 24–26 2003, Lyon, France.

ISO/TC211. Final Draft International Standard 19109 Geographic information - Rules for application schema. Tech. Rep. ISO/FDIS 19109:2003(E), International Organization for Standardization, Technical Committee ISO/TC 211, Geographic information/Geomatics, 2003.

Lemmens, R.L.G. and de Vries, M., 2004, Semantic description of location based web services using an extensible location ontology. In: Proceedings of Münster GI-days, 1-2 July 2004 : Geoinformation and mobility. pp. 261-262. (IfGI prints ; 22).

Lemmens, R.L.G., 2006, Semantic interoperability of distributed geo-services. PhD Thesis, Delft University of Technology, to be published in 2006.

Klien, E. & Probst, F., 2005. Requirements for Geospatial Ontology Engineering. Paper presented at the 8th Conference on Geographic Information Science (AGILE 2005), Estoril, Portugal.

Knippers, R.A. and B.J. Kőbben, 2003, TOP10NL, Model generalisation, phase 1: TOP10NL objects for the scale range of 1:10k to 1:500k. Technical ITC report.

McGuinness, D. L., and van Harmelen, F. OWL Web Ontology Language Overview. W3C recommendation, World Wide Web Consortium, February 2004.

OGC. OpenGIS Geography Markup Language (GML) Implementation Specification Version 3.1.0. OpenGIS Recommendation Paper OGC 03-105r1, Open GIS Consortium, February 2004. Editors: Simon Cox, Paul Daisey, Ron Lake, Clemens Portele and Arliss Whiteside.

Quak, C.W., P.J.M. van Oosterom, M.E. de Vries, B. Bruns and N. Bakker, 2005, TOP10NL in GML3, Geo Info, Volume 2005, Number 11

Stoter, J.E., Generalisation within NMAs in the 21st century, in: Proceedings of the 22nd International Cartographic Conference (ICC 2005), A Coruña, Spain, 9 - 16 July 2005

Stoter, J.E., Generalisation: the gap between research and practice, 8th ICA WORKSHOP on Generalisation and Multiple Representation, A Coruña, July 7-8th, 2005

Schwering, A. and G. Hart., 2004. A Case Study for Semantic Translation of the Water Framework Directive and a Topographic Database. Crete University Press. 7th Conference on Geographic Information Science (AGILE). Heraklion, Greece. Seite(n) 503-510.