Renewal of the base model for geographic information of the Netherlands

Wilko Quak[1], Paul Janssen[2], Marcel Reuvers[3]

[1] Delft University of Technology, OTB, The Netherlands (c.w.quak@tudelft.nl)
[2] Geonovum, The Netherlands (P.Janssen@geonovum.nl)
[3] Geonovum, The Netherlands (M.Reuvers@geonovum.nl)

Abstract
In 2005 a base model for geographic information in the Netherlands was published named NEN 3610. The model consist of a modeling framework (based on the ISO19100 series) and a collection of extensible base classes. In the following years many organizations have built their models in conformance to this base model (IMRO, TOP10NL, IMTOP, IMWA, IMKICH, IMKL). The implementation of the base model strongly enhances the operational aspect of a successful National Geographic Register and related services. Moreover the base model approach has prepared the Netherlands for implementation of a similar approach followed in the INSPIRE initiative. Based on the experiences and national and international developments we are renewing the base model, this will be finished in 2009. In our paper we share our experiences with the base model of the last few years and give an overview of our proposed changes. The changes come from the following three main sources: NEN 3610 users were asked for change proposals, all current NEN 3610 models were analyzed for harmonization opportunities and finally the INSPIRE specifications that will probably influence many datasets under NEN 3610 were checked. Al these proposals were discussed in a NEN 3610 Framework group resulting in a new version of base model. The most notable changes with respect to the old version are:

- The introduction of design patterns as a way to harmonize data models.
- How to handle the fact that all organizations have a different way of looking at the world and hence splitting the same world into different classes.
- Integration of semantic decisions and implementation decisions in one model.

As there are many similarities between the Dutch base model developments and the INSPIRE process we believe that conclusions from this project can be of benefit for the INSPIRE project and vice-versa.

Keywords: INSPIRE, SDI, UML

1 INTRODOUCTION

In this paper we describe the renewal of the NEN 3610 base model for Geo Information in the Netherlands (NEN Normcommisie, 2005). It was the first implementation of ISO 19100 series standards at the national level. At that time few so called domain specific information models were operational or in the process of development. Partly accelerated by the development of base registraties (E-Overheid, 2009) and the increased interest for standardized information exchange through geo information infrastructures, including the National Geo Register (NGR, 2009), many domains expressed their interest in developing NEN 3610 related information models. From three implementations in 2005 over 15 were registered in 2008.

This extensive use of NEN 3610 as a base model for subsequent domain models provided information on shortcomings and implementation bottlenecks. It is experienced that these shortcomings led to different implementations that were not optimal with respect to interoperability. Furthermore, the semantic content of the model was challenged by the increasing number of domains that made use of the base classes. New domains introduced

new requirements on semantics and therefore a new approach on the model principles of NEN 3610.

International developments, especially the data specifications developed in the INSPIRE initiative (INSPIRE, 2007) have led to extensive knowledge of geo-information modeling. NEN 3610 to some extent could implement generic concepts from these INSPIRE data specifications. Practical gain can easily be established by harmonizing NEN 3610 with the INSPIRE data specifications.

The above described reality was behind the decision to start the renewal of NEN 3610 in 2008. Though successful in its way to attract and start implementations of domain models the changed requirements force NEN 3610 to be renewed and evolve from its pioneer role to a more robust corner stone standard.

The rest of the paper that describes the renewal is structured as follows. Paragraph 2 describes the process of how NEN 3610 was renewed. In Paragraph 3 we give an overview of the old NEN 3610 together with an analysis of its shortcomings of that model. Paragraph 4 we propose changes to the original base model and explain how these proposals alleviate the shortcomings. Finally in Paragraph 5 we wrap up.

2 THE RENEWAL PROCESS OF NEN 3610

The NEN 3610 is a national standard and as such registered by the Dutch National Standardization Institute (NEN). Geonovum as the national body for geo-standardization organizes the operational maintenance of this standard. All the public parties that have NEN 3610 domain specific models are organized in the NEN 3610 Framework Group. This group meets periodically and supports harmonization between models but also raises issues for further development of NEN 3610.

Geonovum is the leading party in the project for renewing of NEN 3610. The NEN 3610 Framework Group is closely involved. All members of this Group were asked to formulate their requirements for a new NEN 3610 in so called Change Proposals. As such the requirements of the stakeholders can be found in the combined change proposals. Furthermore an analysis was made of the generic conceptual model of INSPIRE. These information sources together with the existing NEN 3610 served as the main information input for adaptation of the model. In plenary sessions of the NEN 3610 Framework Group this information was combined with a principal discussion on the basic concepts that should underlie the new base model.

In a smaller project group the newly formulated principles were used to establish a proposal for a new base model. After several iterative sessions in the Framework Group the new version is now in the process of being finalized. Arrived at that stage the finalized NEN 3610 will be put forward to the National Standardization Institute. From there on it will be published for a national open consultation and finally after possible adaptations be adopted as a national standard.

3 THE OLD NEN 3610

In this paragraph we give an overview of the base model for geo-information in the Netherlands as it was published in 2005, together with an analysis on whether the model worked in practice as it was intended.

3.1 Basic principles

The main goal of the base model is enhancing the level of interoperability of geo-information exchange in and between specific domains. To this effect NEN 3610 was modeled as a

generic conceptual model of the real world. The real world is modeled in classes that serve as super classes for subsequent domain implementations. Furthermore attributes and enumeration values are declared for which a common use was expected and agreed. As a result the base model provided an independent model of the real world with a strong emphasis on topography.

3.2 Base class

At the top of the hierarchy the base model has one superclass that should be extended directly or indirectly by all classes under the model. In Figure 1 a UML diagram of this base class is drawn:

Figure 1: Base class of NEN 3610

This means that the attributes that are seen in this class are shared by all objects that implement the base model. As this is a heavy burden for domain models nearly all attributes were made optional, only the identification attribute was made obligatory. Many of the attributes are temporal attributes implemention the bitemporal modeling technique for time (Snodgrass, 1999). Many of these attributes are very visible in the model but especially the temporal attributes are not implemented very often.

3.3 A top level hierarcy with semantics

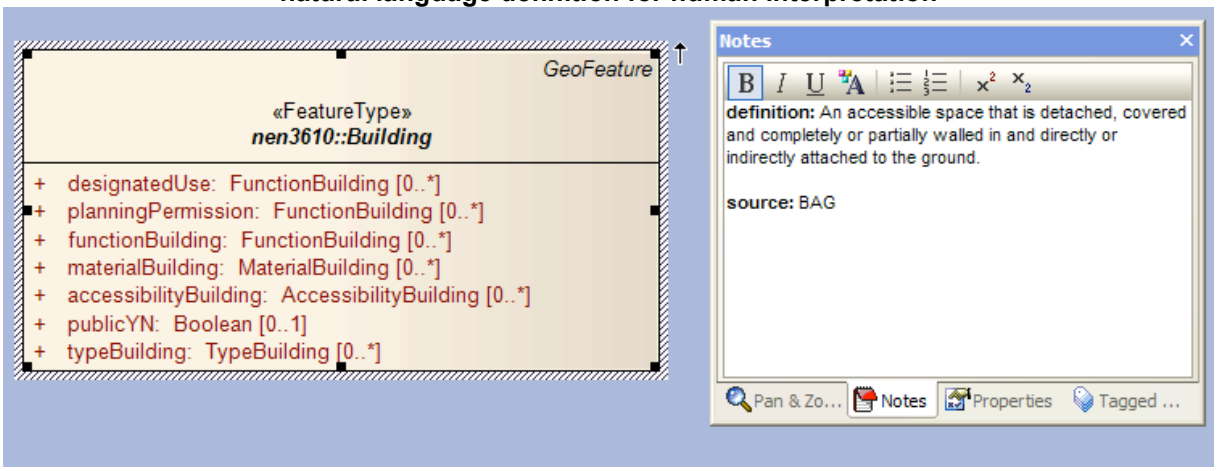
Under the base class GeoFeature a hierarchy of objects with semantics is defined. These meaningful objects mostly referring to things seen in the real word make one of the major differences between NEN 3610 and other models.

Figure 2: Top levels of the hierarchy of the Dutch base model with semantics

3.4 Each class comes with attributes and definitions

Classes in NEN 3610 come with two parts: the UML class diagram and a definition understandable for humans. In Figure 3 we give an example of the class Building with on the left the UML diagram and on the right the formal definition in natural language. Not only all classes have such a definition but also all attributes. The semantics of the attributes in the class diagram follow the UML standard (UML2, 2009). The semantics of the definition for humans is as follows: The definition of a subclass must be a refinement of the definition of a superclass. In order to be made generally applicable all attributes have been made optional: You only use the attribute if it fits your own model. It is not allowed to redefine an attribute in a domain model with different semantics. In practice not everything worked as expected: Very often the attributes as described were only applicable to one domain model and hence did not help in harmonization. Also domains kept on using their own domain specific implementations for concepts that had a generic (but optional) implementation in the base model. The formal definition in natural language did work better, however in some cases the definitions turned out to be ambiguous.

Figure 3: In NEN3610 a class consists of a collection of attributes (left) for computers and a natural language definition for human interpretation



3.5 Making domain models under NEN 3610

During the development of NEN 3610 it was still unclear how to create a domain-model under NEN 3610. It was only stated that a domain model should choose the most appropriate class from NEN 3610 and extend it by subclass. An example of how this was done in TOP10NL (Bakker, 2005) is given in Figure 4.

Figure 4: Implementation of the domain model TOP10NL as an extension by subclass

This idea of extension by subclass did not work in practice for several reasons. Firstly the resulting models UML models became very hard to read. For example, in order to indicate that an optional attribute of the base class is not used in the domain model it had to copied to domain class with cardinality 0 (see the locatie attribute in the above figure) whereas an attribute that is used in the domain model as it is defined in the base model does not appear at all in the domain model. As a result unwanted attributes are much more visible than the regular ones, which makes the models hard to read. Secondly models constructed in the way described above are very hard to automatically convert to a GML application schema. Therefore soon after the adoption of NEN3610 the way of constructing domain models was adapted, now a UML model should implement the intensions of the base model. IMGeo is an example of a model that is made in the new way (IMGeo, 2007).

4 THE NEW NEN 3610

Following the procedure described in Paragraph 2 we are developing a new model that will be finished in 2009. The major changes are described in the subsequent paragraphs. It must be noted that although some changes are big, the impact on existing models will be small.

4.1 Basic principles

The former NEN 3610 focused on harmonization of semantics by way of a comprehensive semantic model with implemented design solutions. However, at this high conceptual level few semantic detail can be modeled that suit the conflicting needs of the different domains. Furthermore it turned out that the principle of one comprehensive model that is extended by domain models is not appropriate.

The new approach is to provide a set of rules that must commonly be applied by domain models in combination with set of semantic base feature types. This less strict way of

harmonization is providing common modeling solutions to common modeling requirements by the way of design patterns.

Furthermore will the NEN 3610 include references to other base models being a model on observation and measurements and a topological (network)model.

The combination of a set of thematic feature types, a base model on measurements and network in combination with common design patterns will provide a harmonized set of building blocks for domain models. By using the same building blocks the domain models will be better harmonized and subsequently interoperability of information exchange will improve.

4.2 References between models

Many of the domain models wish to refer or use objects defined in one of the other domain models, for example the topography model (IMGeo) wishes to include a reference to the Authentic Registration Addresses and Buildings (BAG) for each building. This tricky is because all domain models are developed independently: It is unclear what happens if the definitions of an object that is used in another model changes. The only safe way of referencing is via the base model: Each object has an Identifier that can be used for referencing. Two constructions from INSPIRE are used to solve this problem. We propose a combination of the <<placeholder>> stereotype together with the object referencing design pattern.

4.3 Design patterns

In the new base model we will be using design patterns extensively. In modeling design patterns come from software engineering (Gamma, 1995) where they are defined as a general reusable solution to a commonly occurring problem in software design. Having all the models under the base model implementing the same thing in the same way is a big step towards harmonization. The topics that are handles by the design patterns include:

- How to handle nested attributes.
- In what cases to use the <<union>> stereotype and when to avoid it.
- How to handle null values.
- How to reference from one domain model to another.
- How to model the geometry of a class that is sometimes represented by a point and sometimes by a polygon and other re-occurring problems.
- Etc.

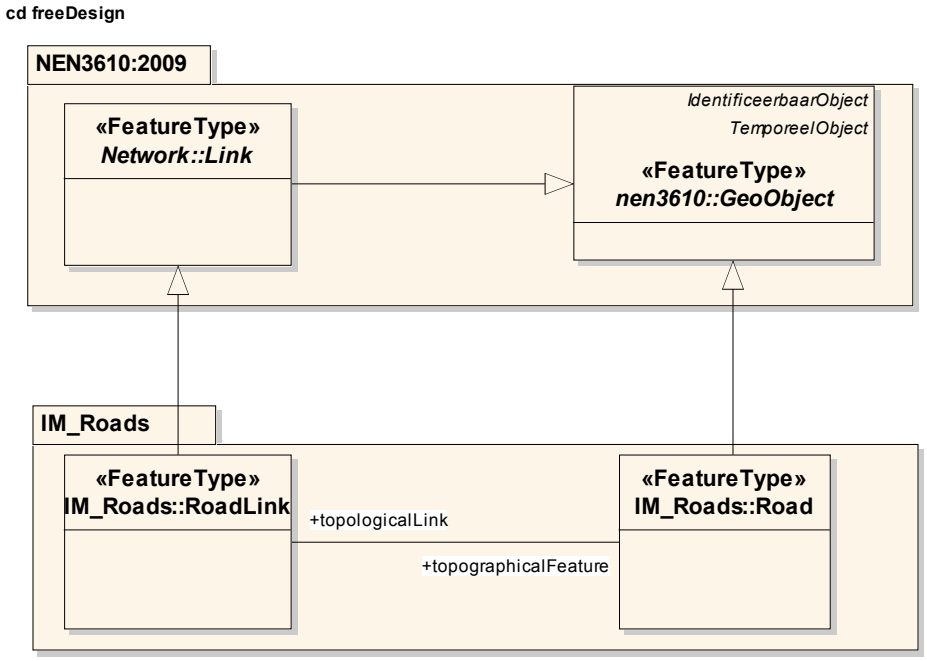
4.4 A semantic model without attributes

One of the much debated parts in NEN 3610 is the class hierarchy under the base class. Many perceived the class names (plus definitions) as a good starting point for their own modeling of the world whereas the attributes as described in the old model (even though they were optional) as not very helpful. Therefore we kept the class names plus definitions in the new model but the attributes are not part of the new model anymore.

4.5 A network model

Modeling of topological network relations is a common issue in several domains. Network of hydrological objects, transportation networks etc. To this purpose a network model is developed that is copied from the INSPIRE generic network model (INSPIRE-D2.5, 2008). The NEN 3610 describes the implementation of this Network model into specific domain models. The use of the network model will lead to a harmonized approach towards modeling in this field.

Figure 5 Example of the use the network model in relation to the NEN 3610. A real world road is represented as a topographical object road as well as a network object roadlink. Each containing their class specific information separately



4.6 A model for Observations & Measurements

Exchange of data about observations and measurements that are related to a specific location is a common practice. The new base model therefore makes reference to an observation and measurement model (O&M). The O&M model is related to the OGC Observation and Measurements standard (Cox, 2007). NEN 3610 describes the relation between the O&M model and the location or geographical feature to which the data are related.

5 CONCLUSIONS

After several years of experience in the Netherlands with a base model for geo information it can be concluded that having a base model helped very much in harmonizing different data-sets within the Netherlands. The model had the following effects on harmonization:

- The process of creating NEN 3610 was done by collaboration between the parties responsible for the domain model. This collaboration was already a first step in the harmonization of the domain models. This collaboration is formalized by having regular meetings between the parties responsible for domain models.
- The NEN 3610 as a base model for implementation in domain models extensively introduced the ISO 19100 series modeling approach in specific domains.
- By describing the different models in a harmonized way (ISO TC/211) was used for modeling the differences between the models were easier to spot and fix. See for example the comparison of TOP10NL and IMGeo (Hofman, 2008).
- The NEN 3610 base model approach did put geo-information models, and more important the specific geo-information domains under a common denominator. In this way it raised a common awareness for common solutions.
- NEN 3610 as part of the Dutch Framework for Geo-information standards combined semantic standards with standards on metadata and services.

Based on our experiences with the old base model we are now renewing the base model, our aim in the new model is to facilitate the harmonization of the underlying models as much as possible:

- In many domain models the same questions arose (how to handle multiple geometries, how to model internal or external references, …). In different domain models different solutions were chosen. By provoding design-patterns for these re-occuring problems we expect the data-models to be more harmonized.
- Within Europe the INSPIRE directive (INSPIRE, 2007) will have a big impact a lot of the domain models within NEN 3610. Where possible we choose to copy the design choices made in the INSPIRE drafting teams and embed them in the new base model.

We expect that with the renewed base model we can even step up the pace of harmonization of spatial data sets in the Netherlands that was started with the publication of the base model in 2005.

REFERENCES

Bakker, N., Bruns, B., Storm, M. (2005) Gegevensmodel TOP10NL (versie 2.3) Kadaster, (In Dutch)

E-Overheid (2009) *Stelsel van basisregistraties* http://www.e-overheid.nl/sites/stelselbasisregistraties last visited: 2009-04-01

Erich Gamma, Richard Helm, Ralph Johnson, John Vlissides (1995) *Design patterns* Addison-Wesley

Hofman A.M. (2008). Integratie TOP10NL-IMGeo, Uitgevoerd in opdracht van: RGI002 'DURP Ondergronden', Technical report, Logica

IMGeo (2007) *Informatiemodel Geografie (IMGeo)* GI-Beraad (In Dutch)

INSPIRE (2007) *DIRECTIVE 2007/2/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community* (INSPIRE)

INSPIRE-D2.5 (2008) *INSPIRE Generic Conceptual Model* Version 3.1

NEN Normcommissie (2005). *Basismodel Geo Informatie* Nederlands Normalisatie-instituut, The Netherlands (In Dutch).

NGR (2009) *Het nationaal georegister* http://www.nationaalgeoregister.nl/ last visited: 2009-04-01

Simon Cox (editor) (2007) *Observations and Measurements – Part 1 - Observation schema* Open Geospatial Consortium Inc Version 1.0

Richard T. Snodgrass (1999). *Developing Time-Oriented Database Applications in SQL* Morgan Kaufmann Series in Data Management Systems

UML2 (2009). *Unified Modeling Language (UML)*, version 2.2. Available at http://www.omg.org/spec/UML/