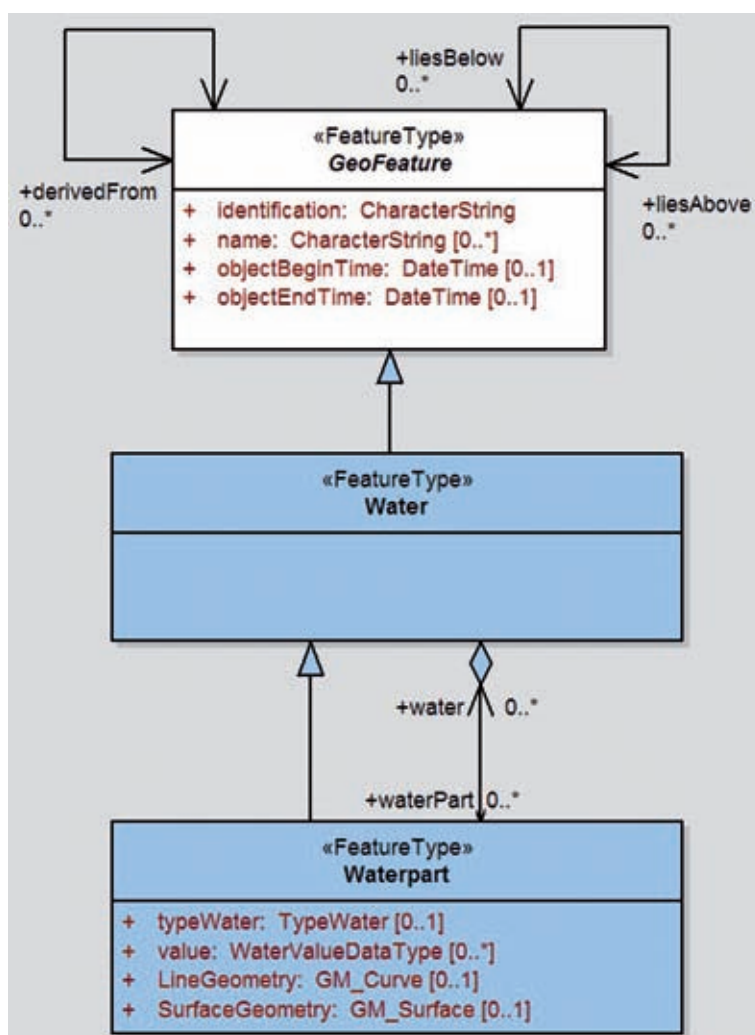


Part 3: Data Model

Standards in Practice

In order for it to be shared, information needs to be defined in such a way that no confusion about its content is possible. Within the geo-sector it is a matter of information about geographic objects, also called features or geo-features. Not only the type of object is recorded, but also its properties or attributes and its relations with other features.

By Huibert-Jan Lekkerkerk



Simple data model for exchanging surface water information.

Data Models

What they are for: Defining information elements, their attributes and their relations with other information elements.

Relevant standards

- ISO 19103: Use of UML for data models
- ISO 19107: Description of geographic and topologic attributes
- ISO 19108: Description of temporal attributes
- INSPIRE: draft implementing rules for data harmonization

Technical implementation

- Universal Modeling Language (UML)

Legal basis: None yet; the first INSPIRE inquiry is currently underway.

A so-called data model is often used to describe information. In everyday practice the use of data models is frequently coupled with relational databases. The difference between a data model and the structure of a database is, however, that the latter is a technical implementation of a previously-defined, logical data model.

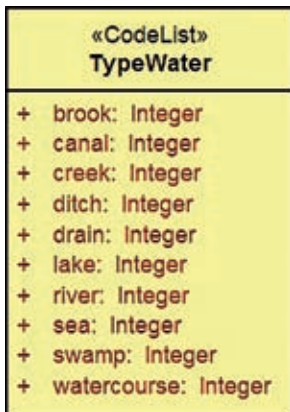
Data Models

Within the 'regular' ICT industry, designing data models has been common for the last decade or so. Every database development warrants the development of a data model. Within the geographic information or ICT industry, the creation of data models was either not possible or simply not performed. A layer structure was usually defined within the software. Specific attributes were then coupled to a specific layer.

With the introduction of object-oriented data acquisition and processing, the storage of information within layers is becoming harder and harder and therefore many organizations are switching to object-oriented data models.

Generic Data Model

If a certain group (or groups) of people can decide on the definition and content of a shared set of objects, a generic data model that makes data exchange between parties possible can be defined. In the Netherlands such a generic geographic data model was developed with NEN, the Dutch national stan-



Code list with water types.

standardization body. The resulting standard, NEN3610: Base model for geography, now serves as an anchor for sector-specific models. Sector-specific

models use the model as the basis from which their own model is extended. Currently over ten different sector-specific models have been developed including the Information Model Water (IMWA), the Information Model Spatial Planning (IMRO) and the Information Model Large-scale Topography (IMGEO).

How Does it Work?

There are international standards for the creation of geographic data models. The main standard used is UML, Universal Modeling Language. This is a very powerful language that can be used to create XML and GML schemas. Using a simple example (see figure) I will now try to explain some of the basics of UML data modeling.

Classes

The basis of any UML model is the objects or classes of which three are pictured in the example (GeoFeature, Water and Waterpart). GeoFeature is a base class of which all other classes are subclasses or specializations (closed arrow). This class contains all the generic properties (attributes), such as identification and name, that are common to the other classes. The class Water is an example of a subclass. This class may seem to have no attributes but in fact inherits all attributes

from the GeoFeature. The class GeoFeature is an abstract class (name is in italics) which is never found in the real world. In order to use the attributes of GeoFeature, one needs to create a specialization such as Water. The class Waterpart inherits in turn all the attributes of GeoFeature (via Water) and adds its own attributes.

Attributes

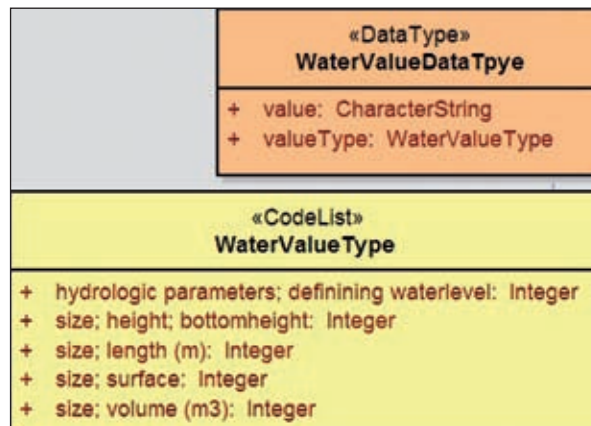
Different types of attributes can be found in the example. There are general attribute types such as ‘text’ [CharacterString] and [DateTime]. There are also specific types for geometries such as PointGeometry [GM_Point] and SurfaceGeometry [GM_Surface]. Further, there are attributes that have a sort of ‘pull down’ list attached. Such a list is called an enumeration list (fixed) or code list (extendible). An example of a code list is given in the figure on the left.

Finally, we can see compound attributes. These are complex attributes that are built around their own attributes. They look like classes but are in fact part of a class. In the example this is the case for the ‘value’ attribute.

Relations and Cardinality

In simply defining classes and attributes, the relationship between classes or the number of times a certain attribute may be used within a class is not defined. With the help of associations, however, the relations between classes can be defined. In the example we see that the class Water is built from Waterparts. This is indicated with an open arrow that may or may not be combined with an open or closed diamond.

Another association is demonstrated with



Compound data type ‘value’ with accompanying CodeList.

GeoFeature which has associations to indicate that it is derived from another GeoFeature or that it is lying above or below another GeoFeature. If, for example, we decide to build a certain water from other waterparts then that specific water changes but is still derived from the original water. The exact time at which the change was made can be indicated in the example with the temporal attributes objectBeginTime and objectEndTime.

Finally the number of times an association or attribute may be used is denoted with cardinality:

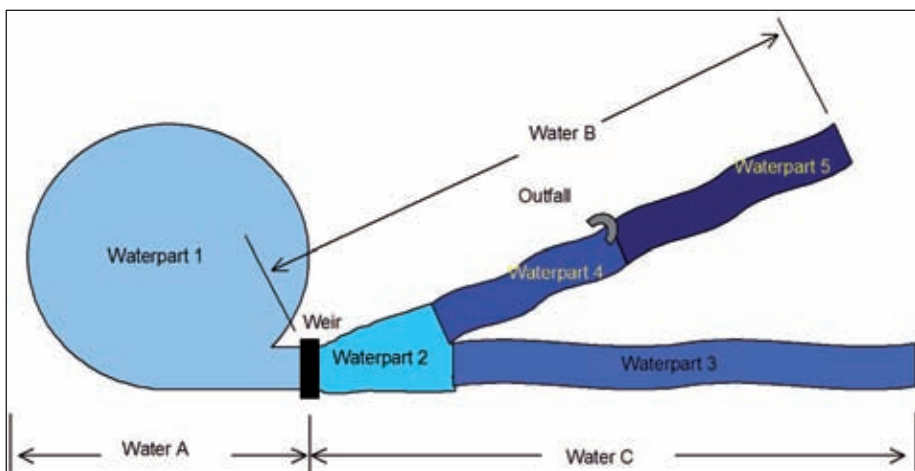
- [1] or no notation: Mandatory attribute or association.
- [0..1]: optional attribute or association. It may be used, but does not have to be.
- [0..*]: optional attribute or association that may be used multiple times.

Legalization

At the moment there is no legal basis for creating data models, although it is common practice. Within the INSPIRE directive, a drafting team is currently investigating whether a common data model such as the Dutch NEN3610 can be developed for Europe.

If such a data model is conceived, it will probably limit itself to the class GeoFeature and will for that class define attributes such as identification as well as temporal attributes.

Huibert-Jan Lekkerkerk (hlekkerkerk@geoinformatics.com) is projectmanager Standards at IDSW and furthermore Editor-in-chief of GeoInformatics.



Building surface waters from water parts.