Reference Information Specifications for Europe (RISE)
Specific Support Activity SSA3-2005-012141
Project co-funded by the European Commission within the 6th Framework Programme (2002-2006)
Integrating and Strengthening the European Research Area – Priority 4: Aeronautics & Space

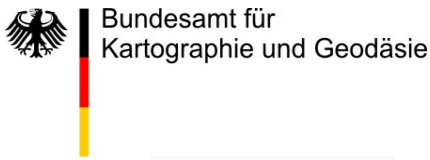
Conceptual Schema in UML
Version 1.1

Deliverable number: 23 [D4.3.1]

Date: 07 September 2007

Project started: 1st September 2005
Duration: 24 months

Project Coordinator: Nick Land, EuroGeographics



Customer:	European Commission
Customer Reference:	Contract FP6-012141
Project Title:	RISE: Reference Information Specifications for Europe
Document Title	Conceptual Schema in UML
Principle Author	Andreas Illert, Sabine Afflerbach
Principle Contributors	Morten Borrebæk, Christine Brönner, Helen Eriksson, Dominique Laurent, Clemens Portele, Lars Stalsberg
Responsible Partner	BKG
Reviewed By	RISE PCMG
Approved by	Nick Land
Filename	RISE23_Conceptual Schema V1.1.doc RISE23_Conceptual Schema V1.1_Annex 1_Matching_table.xls RISE23_Conceptual Schema V1.1_UML ApplicationSchema.doc
Distribution	PU
Contact	http://www.eu-rise.org/ Andreas Illert, Sabine Afflerbach firstname.lastname@bkg.bund.de

This document has been prepared for use by the RISE Project Partners. It can be used or disseminated consistent with the document classification given on the front page (see coding below).	
PU	Public
PP	Restricted to other programme participants (including the Commission Services)
RE	Restricted to a group specified by the consortium (including the Commission Services)
CO	Confidential, only for members of the consortium (including the Commission Services)

Revision History			
Version	Author	Date	Comments
1.0a	Andreas Illert, Sabine Afflerbach	2007-01-09	First Draft
1.0b	Andreas Illert, Sabine Afflerbach	2007-01-12	Comments from consortium members incorporated. Draft version for distribution to the INSPIRE DT Data Specifications
1.0c	Andreas Illert, Sabine Afflerbach	2007-03-09	Recent comments and discussions from the development of Use Cases and Data Product Specification incorporated.
1.0	Wyn Cudlip, Andreas Illert	2007-03-14	Minor changes
1.1	Sabine Afflerbach	2007-09-07	Minor changes to the UML model

Executive Summary

A key outcome from RISE is the definition of a process – a repeatable methodology – for developing, adopting and maintaining harmonised data product specifications. The repeatable methodology is to be tested and improved with use cases. The first step in this process is a high level description of the RISE use cases. Then, an application schema is developed in UML. Finally, the schema will be implemented in a prototype and tested. This document describes the development of the application schema.

The application schema has been designed by a team that was composed of domain experts, GI experts and software engineers. A facilitator from the RISE project managed the process.

For the hydrography theme, RISE could build on guidelines and UML models from the GIS Working Group of the Water Framework Directive. The development of the RISE application schema was fairly straightforward.

For the land cover theme the requirements were less clear than for hydrography. The first draft of the application schema was rejected after discussion. The experts developed an approach based on experience in Norway which should be applicable to the nutrient leakage reporting as required for 2010.

The development of an application schema for elevation suffered from the absence of any documented requirements by the time the work started. The consortium analysed the as-is situation concerning the Digital Elevation Models in the member countries to find applications that relate with the RISE use case. From that analysis resulted in a new use cases, namely the computation of “thalwegs” ("theoretical hydro") in France. Further development was rather delayed but straightforward.

Table of Contents

Executive Summary **4**

Table of Contents **5**

Reference documents **6**

1. Introduction **7**

2. Foundations **8**

2.1. The RISE Use Case 8
2.2. Methodology on Schema Development 8

3. The process towards the RISE Application Schema **10**

3.1. The experts involved, and their roles 10
3.2. Progress in development of application schema: hydrography 12
3.3. Progress in development of application schema: land cover 13
3.4. Progress in development of application schema: elevation 15

4. The RISE Application schema **18**

5. Experience **22**

ANNEX 1 – Matching Table **23**

ANNEX 2 - UML application schema (textual) **23**

Reference documents

The following referenced documents are indispensable for the application of this document.

1. RISE 19 Methodology & Guidelines on Use case & Schema Development. Version 1.1, September 2006..
2. RISE 18 Use Case Document 'Diffuse nutrient leakage reporting to the Water Framework Directive'. Version 1.1, March, 2007.
3. RISE 25 Data Product Specifications
4. ISO/TS 19103:2005, Geographic Information – Conceptual Schema Language
5. ISO/TS 19107:2000, Geographic Information – Spatial Schema
6. ISO 19109:2005, Geographic Information – Rules for Application Schemas
7. ISO 19123:2005, Geographic Information – Schema for coverage geometry and functions
8. ISO/DIS 19131, Geographic Information – Data Product Specification
9. ISO/TS 19139:2006, Geographic information — Metadata — XML schema implementation
10. EUROHARP Description of work, EVK1-CT-2001-00096, December 2001.
11. Panagos, P.and M. van Liederkerke (2004): The Euroharp information system, JRC Publication EUR 21374EN.
12. Vogt, J. (ed.) December 2002: Implementing the GIS elements of the Water Framework Directive. Guidance Document. JRC Publication EUR20544 EN. http://www.ec-gis.org/docs/F2305/GIS-GD.PDF (accessed March 5, 2007)

1. Introduction

The overall aim of RISE is to facilitate the production of data product specifications on the conceptual and implementation level consistent with the relevant international and industrial standards. A key outcome from RISE is, therefore, the definition of a process – a repeatable methodology – for developing, adopting and maintaining data product specifications. The definition of the repeatable methodology in particular addresses issues concerning the harmonisation of heterogeneous data sources.

The repeatable methodology is based on the standards and guidelines laid down by OGC and ISO (ISO 19100 series of standards). The methodology is described in the RISE deliverable 19 "Methodologies and Guidelines on Use case and Schema Development" [1].

The next phase in RISE is testing and improving the methodology with use cases. The first step in this process is a high level description of the RISE use case, to be documented in use case templates and a checklist (see RISE18 "Use Case Document 'Diffuse Nutrient Leakage Reporting to the Water Framework Directive'" [2]). Then, a conceptual schema is developed in UML. Finally, the schema will be implemented in a RISE Test Environment (RTE) and tested.

This document describes activities carried out by RISE to achieve the second step in the process, i.e. the development of the application schema in UML. The scope of the document is

- to describe the actual process that leads to the UML application schema, based on the repeatable methodology that is outlined in RISE deliverable 18 [2]

- to describe the result, a UML application schema for the RISE Use Case 'Diffuse Nutrient leakage reporting to the Water Framework Directive (WFD)'

- to summarise the experience from the process to eventually improve the repeatable methodology

The RISE Data Product Specifications, with clauses specified in ISO 19131, will be defined in a follow-up document: ”RISE25 Data Product Specifications” [3].

2. Foundations

2.1. The RISE Use Case

In its project proposal, the RISE consortium addresses the GMES data harmonisation with a focus on hydrography, land-cover and elevation data themes. During the first phase of the project, relevant GMES projects were analysed. As a result the RISE use case was narrowed to 'Diffuse Nutrient Leakage' reporting to the Water Framework Directive (WFD) [12].

Discussions with domain experts revealed that the data flow in nutrient leakage reporting consists of three major steps:

1. collection and harmonisation of source data for input to the nutrient leakage model
2. computing of nutrient leakage parameters in the nutrient leakage model
3. reporting of nutrient leakage parameters to the EU

The processing of data within the nutrient leakage model is out of scope for RISE. This model has to be considered as a "black box". It appears that different models are set up in the EU member states due to different environmental conditions. A project on harmonisation of the models (EUROHARP) did not come up with a single pan-European solution [10],[11].

As a consequence, the RISE Use Case scenario was narrowed further to the first step in the process described above, i.e. harmonisation of source data to the nutrient leakage model. For prototyping, several scenarios were considered and analysed. It was decided to start with the Swedish example, because the Swedish partner in RISE has good contacts with the Hydrographic Authorities and data was at hand.

The description of the Swedish Use Case was available in July 2006, as a starting point for the development of the Application schema.

2.2. Methodology on Schema Development

The main purpose of the UML application schema in RISE is testing of the repeatable methodology for the development of geospatial data product specifications. This methodology is outlined in the RISE deliverable 19 "Methodology & Guidelines on Use Case & Schema Development" [1].

The methodology consists of five steps:

1. Use Case description

   Result: Use case description, updated glossary

   documented in: Use Case Template [1]

2. Identification of requirements

   Purpose: identify and describe requirements from the use case. These requirements are transformed by a GI expert to a first-cut of the data product specification (first-cut application schema).

   Result: List of feature types, list of requirements

   documented in: Data Harmonisation Checklist [1]

3. As-is-analysis

   Purpose: identifying the information used by the use case actors

   Result: Description of the current situation (per source dataset)

   documented in: Data Harmonisation Checklist [1]

4. Gap Analysis

   Purpose: compares the results of the 'as-is' analysis with the first-cut application schema, evaluates if the identified source material is sufficient to fulfil the requirements in the application schema. Identifies how to extract information from these data sources into the application schema

   Result: Description of data harmonisation issues derived from the identified requirements and taking the as-is analysis into account

   documented in: Data Harmonisation Checklist [1]

5. Application Schema and Data Product Specification

   Purpose: The results from the use case and the analysis are formalised in the data product specification.

   Results:

   1. ISO 19109 application schema specified in UML. A GML application schema is derived automatically from the UML application schema.

   2. Data product specification with clauses specified in ISO 19131. The data product specification includes the application schema in UML and adds the corresponding feature catalogue plus information on data quality, data maintenance, data capture, data delivery, reference systems and metadata.

   3. Updated data dictionary, updated glossary

The steps are not carried out sequentially, but with a considerable overlap to allow for rapid feedback. At every step, potential issues are pushed back to the previous steps to enhance the process, if required (iterative development).

3. The process towards the RISE Application Schema

3.1. The experts involved, and their roles

The development of harmonised data specifications is a complex process. Usually it requires knowledge which cannot be concentrated on a single person. The essential roles in a team are outlined in the deliverable D19 Methodology & Guidelines document [1] as follows:

> Domain expert: provides expertise about the thematic domain and the data to be used in the application
>
> GI architect: provides expertise about geographic information specifications (ISO 19100 series, OGC Implementation Specifications, other GMES/INSPIRE specifications) and information modelling
>
> Software engineer: provides expertise about implementation and deployment aspects of the relevant specifications
>
> Facilitator: manages the process and oversees that all data harmonisation requirements (see the checklist) are identified and adequately addressed

In the RISE Use case, the domain expertise needs to cover the areas hydrography (with a focus on the Water Framework Directive and the nutrient leakage reporting), land cover/ land use (e.g. CORINE) and general reference data (topographic data sets, digital elevation models). With the RISE test areas being located in Sweden, Norway and France, it is important to involve experts that are familiar with the local situation. The experts should be able to communicate in English.

For the development of Application schema in RISE, the following people have been involved in meetings and teleconferences:

Christine Brönner	OGC	facilitator, GI architect
Clemens Portele	OGC	GI architect, software engineer
Jörgen Hartnor	Lantmäteriet	facilitator
Helen Eriksson	Lantmäteriet	facilitator, GI architect
Kerstin Johansson	Lantmäteriet	domain expert (NMCA source data)
Tommy Mukka	Lantmäteriet	domain expert (NMCA source data)
Lars Stalsberg	NVE	Domain expert, GI architect
Astrid Voksø	NVE	Domain expert, software engineer
Daniel Björkert	SMHI	Domain expert, software engineer
Sabine Afflerbach	BKG	GI architect, software engineer
Andreas Illert	BKG	GI architect

Morten Borrebæk	Statens Kartverk	GI architect, domain expert (NMCA source data)
Dominique Laurent	IGN	facilitator (elevation use case), domain expert (NMCA source data)
Ariane Blum	BRGM	domain expert (elevation use case)

BKG	Bundesamt für Kartographie und Geodäsie, Federal Mapping Agency of Germany
BRGM	Bureau de recherches géologiques et minières, Bureau of Geological and Mining Research
IGN	Institut Géographique National, National Mapping Agency of France
Lantmäteriet	National Mapping Agency of Sweden
NVE	Norges vassdrags- og energidirektorat, Norwegian Water Resources and Energy Directorate
OGC	Open Geospatial Consortium
SMHI	Sveriges meteorologiska och hydrologiska institute, Swedish Meteorological and Hydrological Institute
Statens Kartverk	National Mapping Agency of Norway

Six physical meetings were held:

January 31, 2006, Stockholm, Sweden, Swedish RISE Use Case Start-up Meeting

February 28, 2006, Stockholm, Sweden, Swedish Rise Use Case Meeting

August 28, 2006, Arlanda, Sweden, Swedish/Norwegian RISE Use Case Meeting

September 27, 2006, Arlanda, Sweden, Swedish/Norwegian RISE Use Case Meeting

October 19, 2006, Oslo, Norway, Swedish/Norwegian RISE UML Modelling Meeting

November 29, 2006, Orleans, France, Elevation Use case meeting

Besides the Use Case meetings, interviews with Swedish domain experts have been performed with Lantmäteriet in Gävle, the Water Authorities in Stockholm and Kalmar and with SMHI in Norrköping, during a week in March (March 20 – March 24, 2006).

Between the physical meetings, the progress was monitored in teleconferences. Documents were exchanged frequently by e-mail.

3.2. Progress in development of application schema: hydrography

For the hydrography theme, RISE could build on guidelines and UML models from the GIS Working Group of the Water Framework Directive. The development of the RISE application schema was fairly straightforward.

Step 1 – Use case description

By July 6, a first version of the RISE use case description was available, presenting the Swedish use case: 'Diffuse nutrient leakage reporting to the Water Framework Directive in Sweden'. Other material for the RISE application schema includes the Swedish Surface Water Standard (SSWS), the WFD GIS Guidance Document and the EUROHARP Data Exchange Protocol.

Step 2 – Identification of requirements

In a discussion paper (August 17), Andreas analysed the existing information and proposed a first-cut application schema based on the EUROHARP model. Reasoning: EUROHARP already integrates hydrography and land cover, it is tailored to the RISE use case of nutrient leakage reporting, and it already provides a schema described in UML. At the Swedish/Norwegian RISE Use Case Meeting in Arlanda, August 28, the participants raised concerns that the requirements from EUROHARP do not coincide with the WFD. It was decided that the hydrography requirements should rather be taken from the WFD. Even though EUROHARP should be compliant with the WFD, the WFD is still the original source.

Step 3 – As-is analysis, and Step 4 – Gap analysis.

Sabine prepared a matching table (September 21) that relates the feature types and attributes from EUROHARP, WFD, Sweden "Red Map", Sweden SMHI, Norwegian test data (provided by Lars Stalsberg) and additional hydrographical classes from the Norwegian text document "Innsjøer og vassdrag" sent by Morten on September 14, 2006.

Step 1 – Use Case description, through to step 4 – Gap analysis

By September 22, Christine provided the description of the Norwegian use case 'Hydrography input data processing along the border to the Sweden'. The use case description feeds into steps 3 – As-is analysis, 4 – Gap analysis and 5 – Harmonisation approach. The matching table was updated in the process. Another Swedish/Norwegian RISE Use Case Meeting was held in Arlanda, September 27, where the matching table was discussed and decision on test datasets was made. The participants of the meeting identified differences in terminology between Norway and Sweden.

Step 5 – Harmonisation

Lars provided a proposal on harmonised terminology (September 29). At a UML modelling meeting in Oslo, a draft schema was discussed in detail including the relationships with land cover, and put into proper UML.

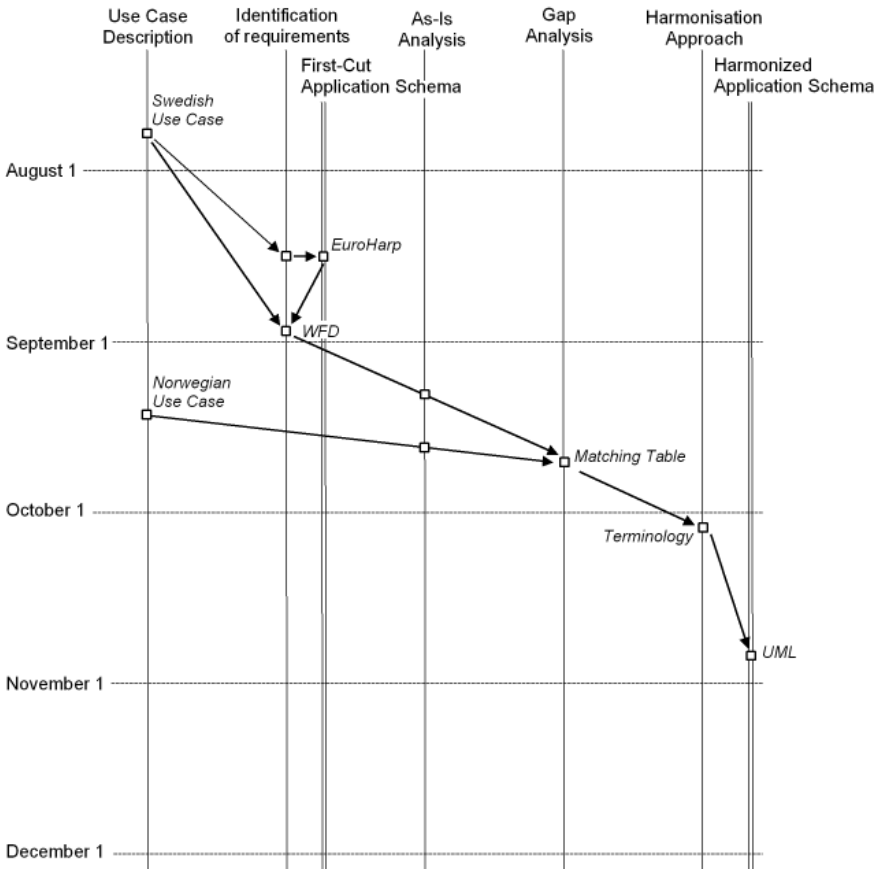


Figure 1: Sequence model for the development of RISE application schema – hydrography

3.3. Progress in development of application schema: land cover

For land cover the user requirements were less clear than for hydrography. The first-cut application schema was rejected after discussion. The use case scenarios for Sweden and Norway describe the 2005 reporting, they do not deal with the processing of land cover data for the WFD. Finally, the experts developed an approach based on experience in Norway which should be applicable to the nutrient leakage reporting as required for 2010.

Step 1 – Use case description

By July 6, a first version of the RISE use case description was available, presenting the Swedish use case: 'Diffuse nutrient leakage reporting to the Water Framework Directive in Sweden' based on the Swedish Nutrient Leakage Model TRK. Other material for the RISE land cover application schema includes specifications for CORINE, the Swedish Red Map and the EUROHARP Data Exchange Protocol. The WFD has not stated any requirements on land use / land cover.

Step 2 – Identification of requirements

In a discussion paper (August 17), Andreas analysed the existing information and proposed a first-cut application schema based on the EUROHARP model, for the same reasons as described under hydrography. EUROHARP provides a predefined list of 15 land cover types, plus a detailed list of crop types that was not considered for the first-cut application schema. At the Swedish/Norwegian RISE use case meeting in Arlanda, August 28, the participants discussed whether RISE should use the requirements from EUROHARP for the land use themes or not. The EUROHARP model was considered far too detailed for the RISE use cases. Decision: compare requirements from the Swedish TRK model and the Norwegian

Teotil model with the requirements from EUROHARP. At a teleconference one week later, there was a discussion on what Swedish test data should be used for land use: the classes that result from the current WFD reporting procedure described in the use case report, or the Swedish CORINE. No decision was made about this.

Step 1 – Use Case description

Morten provided a description of data used in Norway ("Markslag" September 14). Christine provided the description of the Norwegian use case (September 22) elaborated by Lars Stalsberg (NVE).

Step 3 – As-is analysis, and Step 4 – Gap analysis.

Sabine prepared a matching table (September 21) that relates the feature types and attributes from CORINE Sweden, TRK Sweden, Red Map Sweden, Teotil Norway, EUROHARP and land cover classes from the Norwegian document "Markslag" sent by Morten on September 14.

Step 2 – Identification of requirements, leading to 5 – Harmonisation

At the Swedish/Norwegian RISE use case meeting in Arlanda, September 27, a new approach was introduced for the classification of land cover. The approach follows the Norwegian example where land cover types are structured according to their effect on nutrient leakage. Adapting that approach to RISE, the land cover classes from Swedish CORINE were aggregated, based on the idea that these groups would have the same coefficient with respect to the diffuse nutrient leakage. The RISE classes use the same naming as CORINE. The participants decided on test data: Sweden will use the Swedish CORINE (SMD) dataset, together with the clear-cut database. Norway will use the Markslag dataset.

The new approach to classification of land cover is not reflected in the current Swedish and Norwegian use case scenarios, which refer to the 2005 reporting procedure. For a consistent documentation of the RISE process it would have been required to re-iterate and set up a use case description for the new approach, referring to a (proposed) 2010 reporting procedure.

Step 5 – Harmonisation

After consultation of Norwegian domain experts on diffuse nutrient leakage Lars introduced some changes on the RISE land cover classes (October 3).

At a UML modelling meeting in Oslo (October 19), the schema was discussed in detail including the relationships with hydrography, and put into proper UML.

At a teleconference on December 6, inconsistencies in terminology were identified and discussed. A major issue was 'land use' versus 'land cover'. The participants preferred the term 'land cover' to 'land use' because the RISE classification is built on CORINE. This had some effects on the nomenclature in the RISE schema.

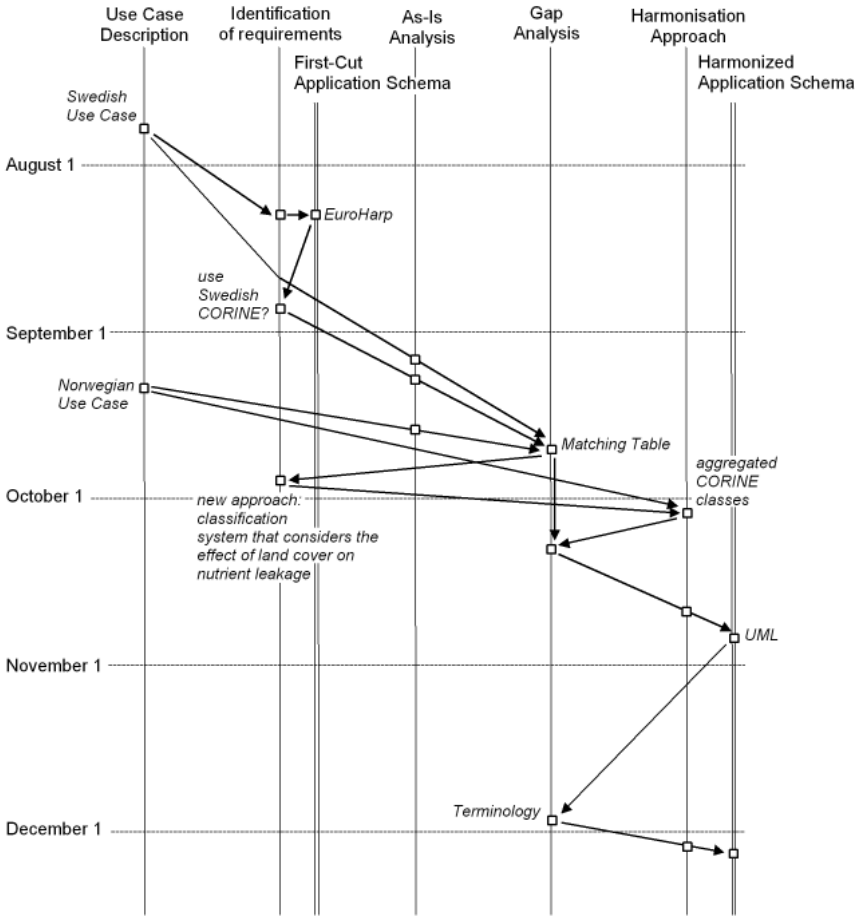


Figure 2: Sequence model for the development of RISE application schema – land cover

3.4. Progress in development of application schema: elevation

The development of an application schema for elevation suffered from the absence of any documented user requirements by the time the work started. Initial use case scenarios were derived from the current practise of WFD data processing in Norway and Sweden, but rejected as they did not contribute to the harmonisation of source data i.e. digital elevation model (DEM). The consortium then analysed the as-is situation concerning the DEM products in the RISE countries to find applications that relate with the RISE use case. From that analysis resulted a new use case, namely the computation of thalwegs ("theoretical hydro") in France. Further development was rather delayed but straightforward.

Step 1 – Use case description

By July 6, a first version of the RISE use case description was available, presenting the Swedish use case: 'Diffuse nutrient leakage reporting to the Water Framework Directive in Sweden'.

Step 2 – Identification of requirements

In a discussion paper (August 17), Andreas analysed the existing information. The Swedish Use case introduces an attribute 'mean height values in meters' to the subcatchment features which is derived from national DEM data. The use of this attribute was not confirmed by other partners of the Swedish consortium working on the WFD. The WFD itself has not stated any requirements for elevation. The requirements from EUROHARP were considered rather generic and not really useful.

The consortium discussed the situation. The partners agreed that RISE should not simplify the issue to the computation of an attribute, but consider an approach towards the harmonisation of elevation source data which means Digital Elevation Model (DEM).

Step 3 – As-is analysis

Information on national DEM data was provided from Norway (September 14) and Sweden (September 18).

Step 2 – Identification of requirements

Christine analysed the role of elevation in Swedish and Norwegian use case models. Conclusion: the elevation theme is represented with attributes ”mean height” (Sweden) and ”slope” (Norway), but no clearly traceable requirements for nutrient leakage reporting could be associated with these attributes (September 29).

Step 1 – Use case description

A teleconference at the 6th October concludes with a proposal for additional use cases.
1. from GMES: pan-European DEM for the rectification of ortho-imagery in the landmonitoring fast track service.
2. BRGM will develop a use case on WFD elevation requirement

Step 2 – Identification of requirements

Andreas outlined the DEM requirements from GMES landmonitoring fast track service (October 13)

Step 3 – As-is analysis

Information on national DEM was provided by France (October 19) and Sweden (November 13)

Step 1 – Use case description

At a meeting with BRGM on the 29th of November, a use case was found that is relevant for RISE: computation of IDPR (Indice Développement Persistence des Réseaux). This IDPR is computed by comparing the network of thalwegs or "theoretical hydro" (deduced from the DTM) with the network of "real hydro".

Step 2 – Identification of requirements, leading to Step 5 – Harmonisation

It appeared that the requirements from the BRGM use case and the GMES use case almost coincide. Both applications need DEM data at a similar level of resolution and accuracy. This requirement is met by the national DEMs used. As a result, the BRGM use case was adopted by RISE and incorporated to the UML schema and the data product specification.

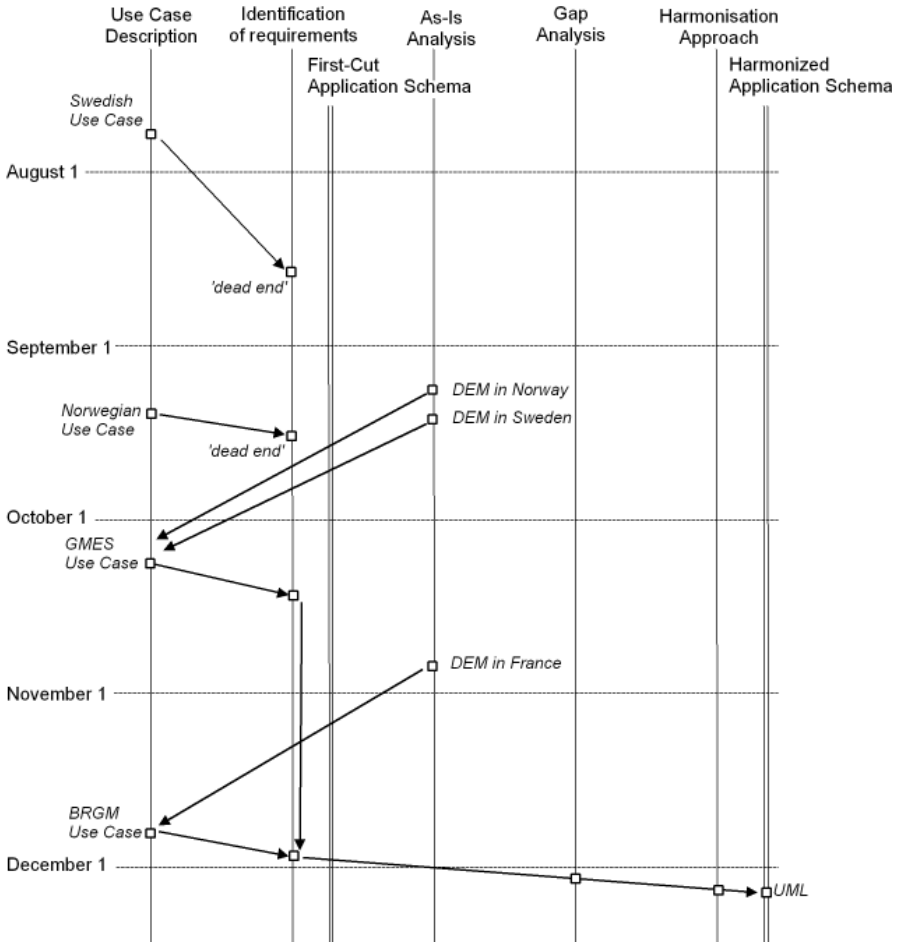


Figure 3: Sequence model for the development of RISE application schema – elevation

4. The RISE Application schema

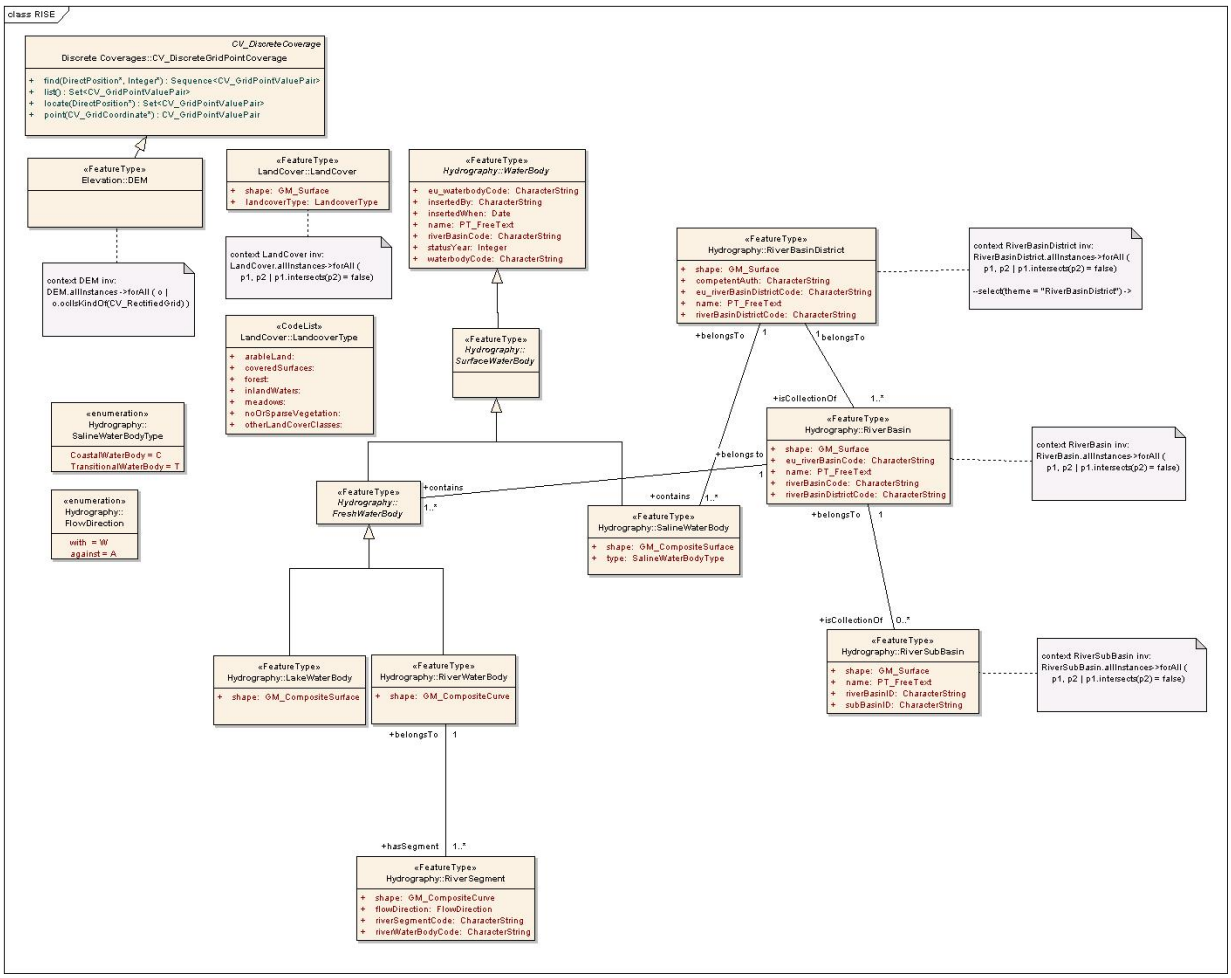


Figure 4: RISE UML Model

For the UML modelling the RISE team used the software Enterprise Architect, version 6.5 from Sparx Systems that is based on UML 2.1.

RISE took advantage of the existing UML model of the Water Framework Directive (see Figure 5) and adapted it to the RISE requirements (see Figure 4).

A textual description of the UML is provided in Annex 2 to this document.

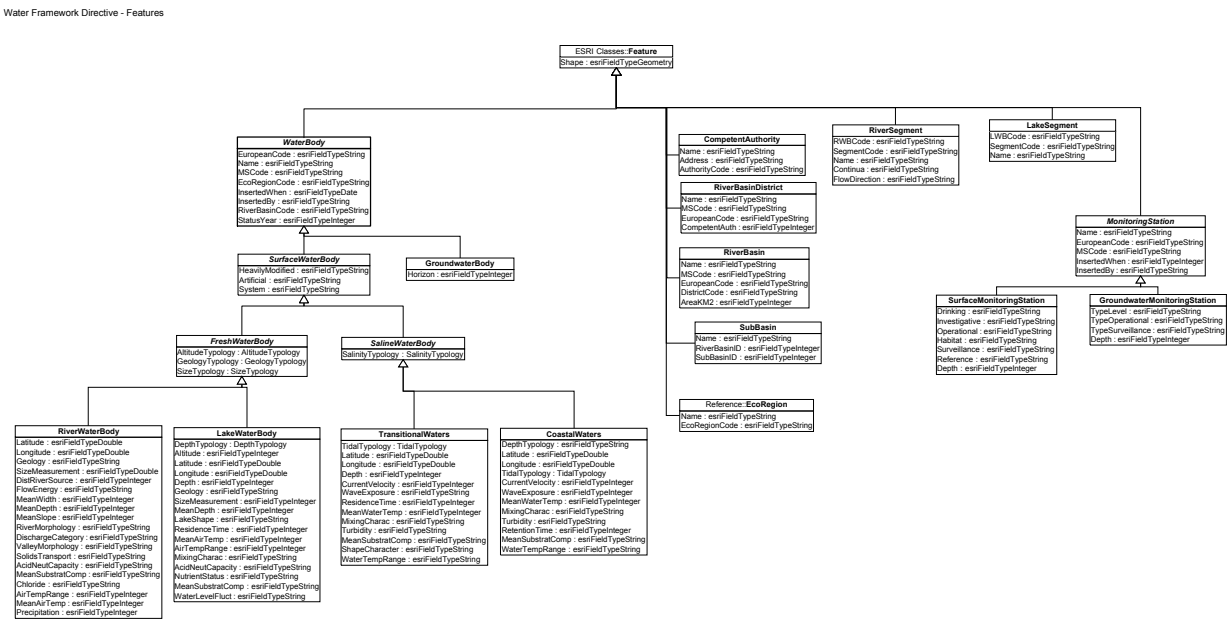


Figure 5: Original UML model of the Water Framework Directive

In a first step, RISE added two feature type classes, one for landcover and one for elevation named 'DEM' (see Figure 6). The DEM feature class inherits from the class CV_DiscreteGridPointCoverage of ISO 19123 "Schema for coverage geometry and functions".

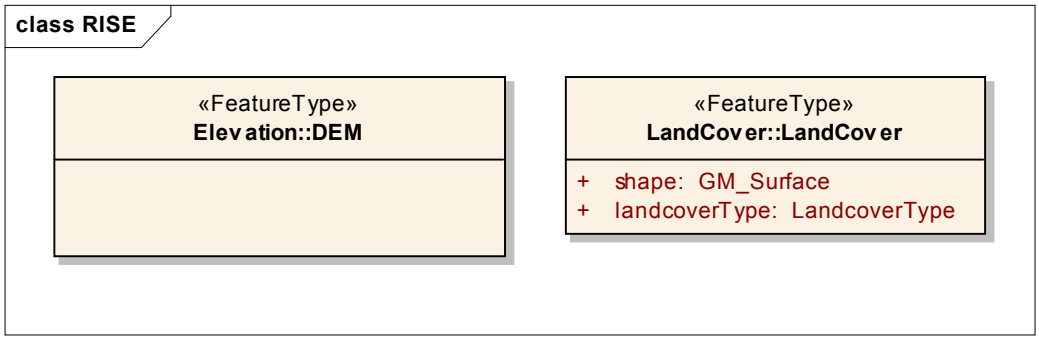


Figure 6: Extract from the UML Model showing Feature Type DEM and Feature Type LandCover

Additionally the DEM is defined as type of the ISO class CV_RectifiedGrid. In the ISO standard 19123 "Schema for coverage geometry and functions" the CV_RectifiedGrid is defined as follows: "A rectified grid shall be defined by an origin in an external coordinate reference system, and a set of offset vectors that specify the direction and distance between the grid lines within that external coordinate reference system. … If the spacing is uniform, then there is an affine relationship between the grid and external coordinate system, and the grid is called a rectified grid. If, in addition, the external coordinate reference system is related to the earth by a datum, the grid is a georectified grid." This definition is set via a constraint written in OCL (Object Constraint Language), see Figure 7.

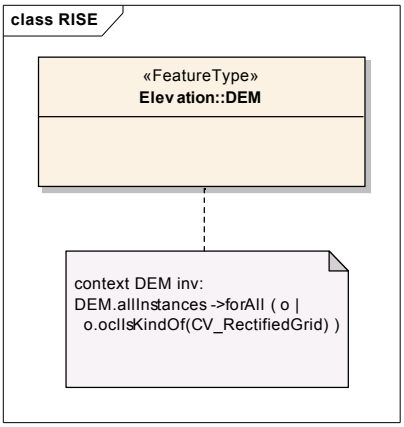


Figure 7: Constraint of the feature type DEM

For the vector data, RISE uses GM_CompositeCurve for lines and GM_Surface/ GM_CompositeSurface for areas as defined in ISO standard 19107 "Geographic information – Spatial schema" to ensure network topology. The GM_CompositeCurve is defined as follows: "A composite curve (GM_CompositeCurve) shall be a GM_Complex with all the geometric properties of a curve. These properties are instantiated in the operation "curve". Essentially, a composite curve is a list of orientable curves (GM_OrientableCurve) agreeing in orientation in a manner such that each curve (except the first) begins where the previous one ends." Correspondingly GM_CompositeSurface is defined as "a GM_Complex with all the geometric properties of a surface, and thus can be considered as a type of orientable surface (GM_OrientableSurface). Essentially, a composite surface is a collection of oriented surfaces that join in pairs on common boundary curves and which, when considered as a whole, form a single surface."

For the landcover FeatureType, RISE attaches a codelist where the seven predefined land cover types are listed as attribute values. An alternative would have been to introduce seven individual landcover classes but during the discussions this was considered to be too complex.

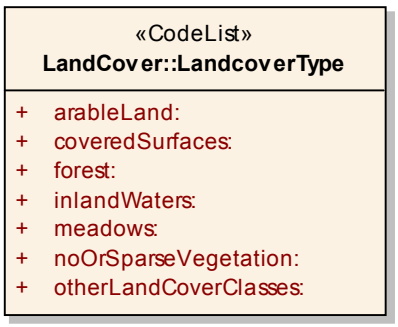


Figure 8: Code list LandcoverType

In order to be ISO compliant RISE shall use data types defined in the ISO standard 19103 "Conceptual schema language", e.g. “CharacterString”. These data types have to be imported in the UML document to be able to use them properly.

One of the RISE requirements is multilingualism. In order to store one geographic name in different languages at one feature, the attribute "name" for the features "WaterBody", "RiverBasinDistrict", "RiverBasin" and "RiverSubBasin" is defined with the data type PT_FreeText. This data type is defined in ISO 19139 "Geographic information - Metadata – XML schema implementation" as a multi-language free text data type.

In the original WFD model several feature types carry attributes called MsCode (unique code for the water body defined in the Member State) and EuropeanCode (unique identifier at European level, including the 2 character ISO Country Code plus the unique code in the member state). The same name of attribute type is used for different coding systems. To avoid these ambiguities, RISE introduces different attribute names. The model now contains attributes called "riverBasinDistrictCode", "eu_riverBasinDistrictCode", "riverBasinCode", "eu_riverBasinCode".

Some attributes only allow certain values. For example the attribute flowdirection only accepts the character W = with or A = against. To express this restriction RISE introduced an enumeration called Flowdirection with these two items (see Figure 9).

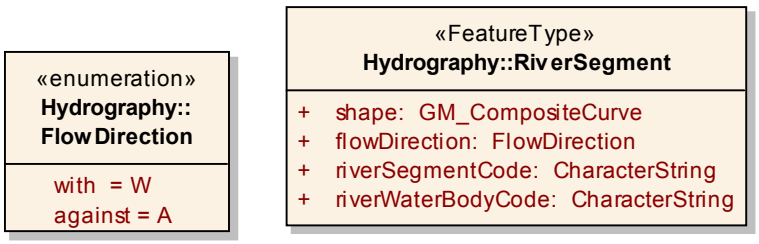


Figure 9: Enumeration FlowDirection

5. Experience

Proper identification of requirements turned out crucial to the development of application schema. Once the requirements were identified and accepted, the harmonisation approach could be developed quite fast. In the opposite, when requirements were not clearly extractable much time and efforts had to be invested in the iterative approach to consolidate the use case.

In the guidelines document, the single entry point to the development of application schema is the use case description, with the requirements to be derived from that document. In reality the conditions were found to be different. For instance, the requirements for the land cover theme could not be extracted from the use case document, because the RISE use case refers to the WFD reporting as it will be in 2010 while the use case scenarios describe the procedure as it is for the 2005 reporting. The domain experts finally developed an approach that reclassifies CORINE classes according to their P and N characteristics. This approach is yet fictitious and would need to be elaborated to a proper use case. If no user requirements are known, then another entry point to the development of application schema might be the analysis of the existing data sources, with the first-cut application schema to be derived as the least common denominator. A third entry point to the development of application schema is the adoption of existing harmonised models. This happened with the hydrography theme, where RISE adopted the WFD model and adapted it to the nutrient leakage use case.

The matching table turned out the most important document in the preparation of the conceptual model. In this role it complements the checklist by describing and relating the semantic models in very detail at the level of feature types and attribute values. From experience we recommend to use the checklist document for overview purpose only, while details are kept and maintained in supporting documents such as the matching table, to be referenced from the checklist document. Likewise the checklist, the supporting documents should be continuously evaluated and improved during the spiral process.

RISE follows the ISO 19100 guidelines in recommending formal modelling languages. However, it should be noted that the essential working documents within the team of experts had been simple files in MS-Word or MS-Excel structure. The model was described in UML only after the group had a clear idea on the content and structure of the data.

In comparing the national data models, the language posed a significant obstacle. Even when national and English terms appear to coincide, it required a native language speaker with domain expertise to confirm (or reject) the proposed match.

According to the RISE deliverable 19 "Methodologies and Guidelines on Use case and Schema Development" [1], the methodology should result in an application schema specified in UML, a data product specification with clauses specified in ISO 19131, an updated data dictionary and an updated glossary. The data dictionary and the glossary were not treated in the tests described above. RISE has used dictionaries and glossaries from the Water Framework Directive and CORINE, but no mechanisms were available to feed back the (minor) changes and amendments to these repositories.

ANNEX 1 – Matching Table

The Matching Table used to compare the existing data descriptions to the required harmonised data specifications is provided as a separate Excel spreadsheet called: RISE23_Conceptual SchemaV1.1_Annex1_Matching_table.xls

This document relates the feature types and attributes from various sources (e.g. EUROHARP, WFD model, Sweden "Red Map") with the RISE application schema.

This is the result of the step ‘gap analysis’ and it complements the information provided in the Checklist.

ANNEX 2 - UML application schema (textual)

See separate document: RISE23_Conceptual SchemaV1.1_UMLApplicationSchema.doc

This document presents the UML schema in textual description. It has been derived automatically with the Enterprise Architect tool.