TESTING CARTOGRAPHIC DATA HARMONIZATION IN THE INSPIRE CONTEXT: THE CASE FOR HYDROGRAPHY

Rui Reis[1]; Dolors Barrot[2]

[1] Instituto Geográfico Português – Rua Artilharia Um, 107, 1099-052 Lisboa
[2] Institut Cartogràfic de Catalunya – Parc de Montjuïc, 08038 Barcelona

(rui.reis@igeo.pt; dolors.barrot@icc.cat)

ABSTRACT

This paper presents the approach followed in the GIS4EU project for testing the draft INSPIRE data specifications and presents the results achieved for the theme hydrography. The objective of the project is to provide base cartography datasets in four themes for Europe, ensuring cross-scale, cross-language and cross-border interoperability and accessibility. The themes used in the project are administrative units, hydrography, transportation networks and elevation.

The call for participation in the testing phase of the draft data specifications for themes of annex I of the INSPIRE directive admitted two kinds of tests. namely, transformation testing and application testing. In this paper we present the methodology and the results of the transformation from the data models of the datasets available in the project to the INSPIRE data model and more specifically for the theme hydrography but the same approach was used in the project for the themes Administrative Units and Transportation Networks.

The core of the testing consisted on the comparison of the data models of the eleven datasets of the project, from nine producers, with the INSPIRE draft data specifications and the corresponding identification of features and attributes in common on the different data models for the theme hydrography. This process was complemented with a critical analysis of the process in order to identify features not present in the INSPIRE data model that the data producers think are needed, the identification of elements in the INSPIRE data model that are not relevant and the identification of the problems found during this process.

1. Introduction

The objective of the GIS4EU project is to harmonize datasets from different producers across Europe in four themes (administrative units, hydrography, transportation networks and elevation) and to ensure cross-scale, cross-language and cross-border interoperability and accessibility. This will be accomplished complying with international standards and to the requirements of the INSPIRE directive.

The partners providing data to the project are from five different countries from the local, regional and national levels. These data producers work with research and technological partners from another five countries and a GI users association.

GIS4EU puts the emphasis on the data and thus great importance is given to the datasets available for the project. Thus, one of the first steps was to document the data in order to formalise the data models needed in the subsequent phases of the project.

The objective is to create common data models, one for each theme, and to develop rules and guidelines of harmonization, aggregation and exposition of the project datasets.

The establishment of common data models is done taking into account the INSPIRE results, namely the work being developed by the data specifications drafting team towards the development of implementing rules. It was in this context that the project participated in the testing phase of the draft data specifications for themes of annex I of the INSPIRE directive.

In this paper we present the approach and results of the participation in the testing of the draft INSPIRE data specifications for hydrography using the data available in the project, namely:

- local data at large scale: provided by the Magistrato Alle Acque di Venezia (MAV) and Comune di Genova (CGE) at scales 1/1000, 1/2000 and 1/5000.
- regional data produced according to the DBPrior10k Italian specifications: provided by Regione Liguria (RLIG), Regione Piemonte/CSI (RPIE), Regione del Veneto (RVEN) and Regione Friuli Venezia Giulia (INSIEL) at the 1/10000 scale.
- data used for reporting in the context of the Water Framework Directive: provided by the Research Institute of Geodesy and Cartography in Bratislava (SK50) at 1/50000 scale.
- general purpose cartographic data: provided by Institut Cartogràfic de Catalunya, at scales 1/5000 (BT5M) and 1/50000 (BT50M), and Instituto Geográfico Português, at scales 1/250000 (ERM_PT) and 1/1000000 (EGM_PT).

2. The process for the definition of implementing rules on interoperability in INSPIRE

The implementation rule on interoperability of data sets and services nowadays is a set of specifications referred to the themes of Annex I of the INSPIRE Directive. Its development has been done in two steps:

- conceptual framework and specification methodology
- data specifications of each theme

More than 40 European experts have been involved in the process, organised in thematic working groups (TWGs). Each one was composed by 4-6 GIS/IT experts, 1 facilitator, 1 editor and 1 JRC member.

The process was leaded by the Data Specifications Drafting Team (DT-DS) who provided the conceptual framework through the following documents: Definition of Annex Themes and Scope [DS-D2.3], Generic Conceptual Model [DS-D2.5], Methodology for the development of data specifications [DS-D2.6] and Guidelines for the encoding of spatial data [DS-D2.7].

The schema in figure 1 shows the process followed by the TWGs according to D2.6 to develop the data specifications for each theme.

It is an iterative process; initially a detailed description of the application schema and feature catalogue has been developed taking into account the previous documents, the reference material provided by SDIC/LMO, the users requirements coming from a survey and the use cases identified and described regarding requirements for the data model.

The first drafts of data specifications were reviewed by the TWG and DS-DT and the updated version of the application schemas and feature catalogues according to the internal comments resolution were produced including thematic metadata and quality element.

Afterwards, the period of implementation, testing and validation was opened ending the process with the new theme data specifications taking into account the comments and testing reports.

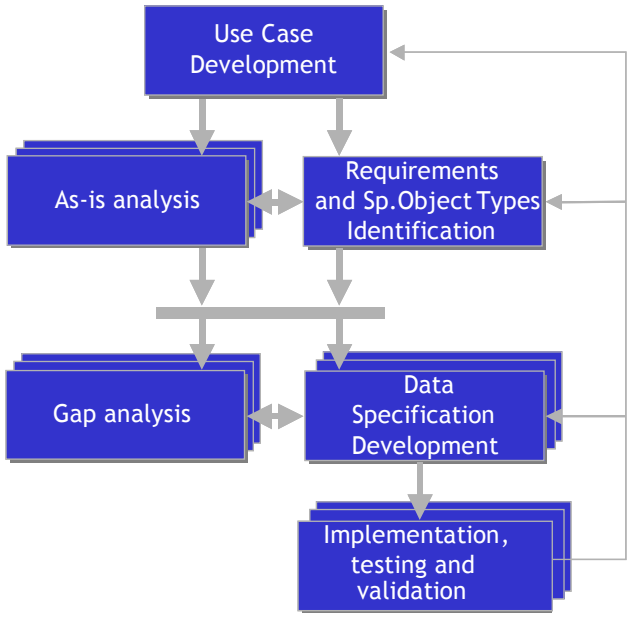

Figure 1 – Methodology recommended by DT-DS for developing INSPIRE data specifications (source: D2.6 v2.0)

3. INSPIRE data model for hydrography

Before starting with the hydrography model it is convenient to set up the theme scope which was established as it follows: Hydrography theme covers all inland water and marine areas covered by river basin districts as defined by the Water Framework Directive (Directive 2000/60/EC), excluding groundwater because is treated under the Geology theme in the Annex II.

The INSPIRE Hydrography data model in Unified Modelling Language (UML) is built around the three following use cases:

- Spatial analysis and modelling: GIS techniques are essential for the derivation of information layers for water management and planning policies and activities (characteristics of water bodies and water ecological and chemical status) based on a hydrographic network.
- Mapping: hydrography is a basic reference layer for other subjects; it includes the representation of main physical waters and related objects.
- Reporting: the implementation of the WFD or other European directives requires the handling of spatial data for reporting to the Commission about quality (particles, pollutants etc.) and quantity of water. Hydrography theme will include the reporting units although the reporting matters will be modelled in the annex III themes as Environmental monitoring facilities, Area management/restriction/regulation zones and reporting units.

Bearing in mind the scope it was necessary to add to the model some features, called "placeholders", for features of other INSPIRE themes but in line with the developed use cases.

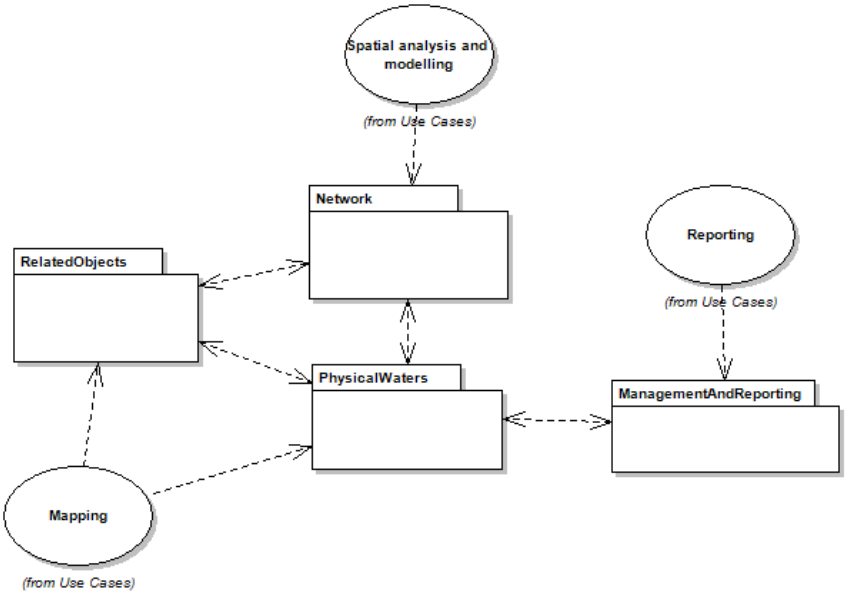

Figure 2 – Hydrography data model and use cases relations (source: INSPIRE portal)

2.1 Spatial analysis and modelling

There are a great deal of studies and works at different resolution levels to prevent flooding risk, to maintain water ecosystems, to determine the impact of water contaminations etc. based on spatial analysis requiring a hydrographic network. Elements in networks are

handled as nodes, links, aggregated links (to define routes) and areas. These abstract classes are extended, reused and specialised in various feature types within the hydrography network data model in the same way as it has been done in the Transport Networks data model.

Moreover, most of these scenarios need elements as pumping or monitoring stations, pipes bridges, dam etc., which can not be properly considered hydrographic elements. They are included in the "RelatedObject package" as placeholders and are identified in the network as a Constriction.

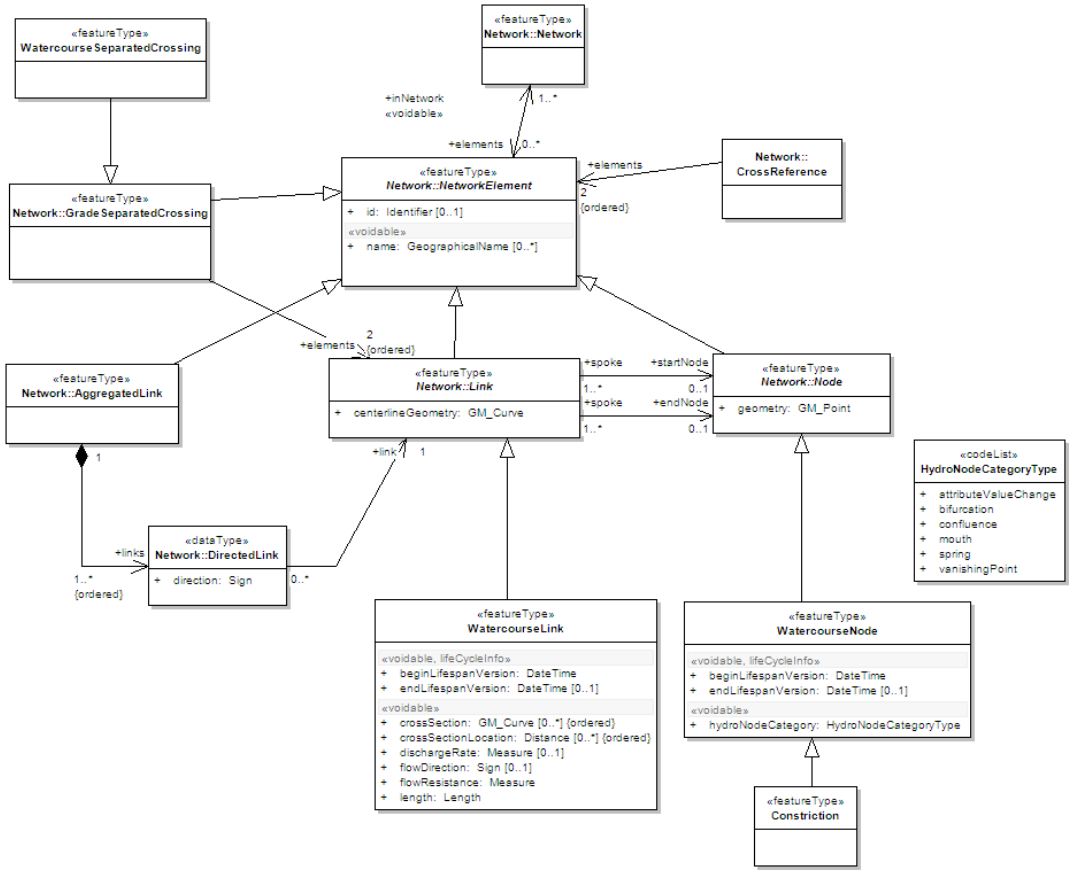

Figure 3 – Hydrography data model: Network package (source: INSPIRE portal)

2.2 Mapping

The main hydrographic elements in additions to roads, administrative units and geographical names can provide a map background to understand place relationships and for orientation.

There are also elements not relevant for spatial analysis or reporting as rapids or falls but very helpful to be located which are included in the “RelatedObject package”.

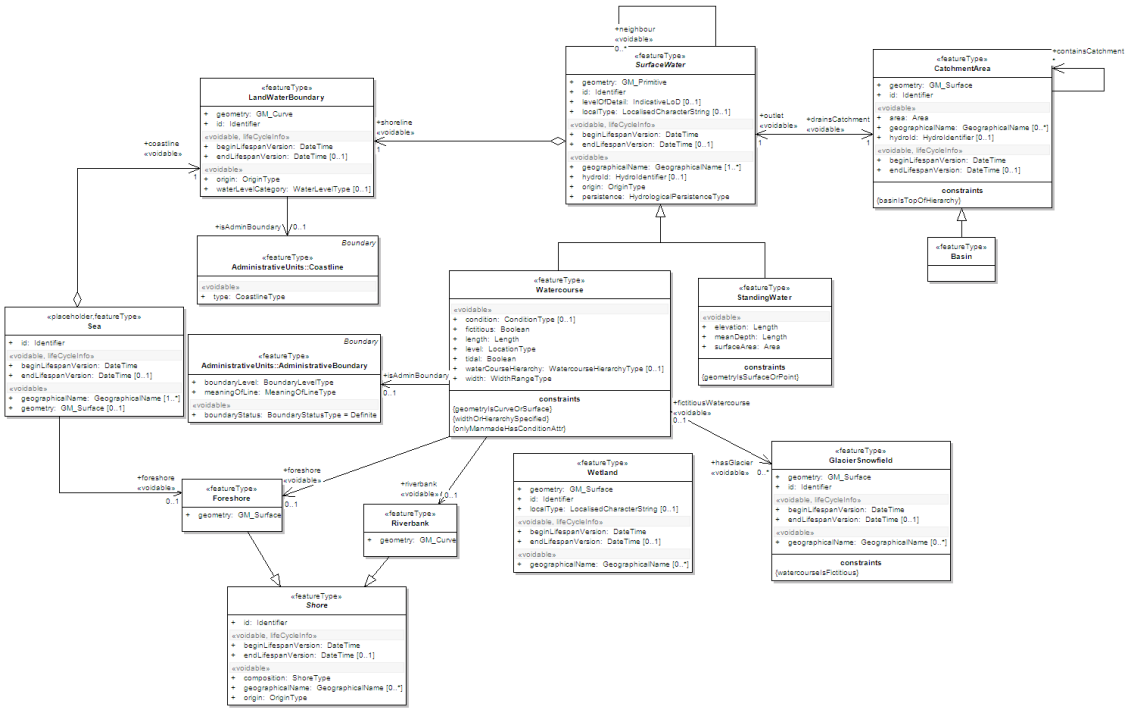

Figure 4 – Hydrography data model: PhysicalWaters package (v2.0) (source: INSPIRE portal)

2.3 Reporting

The reporting use case has been focused on the WFD requirements for mapping and reporting; features definitions came from the WFD and some attributes are included according to it.

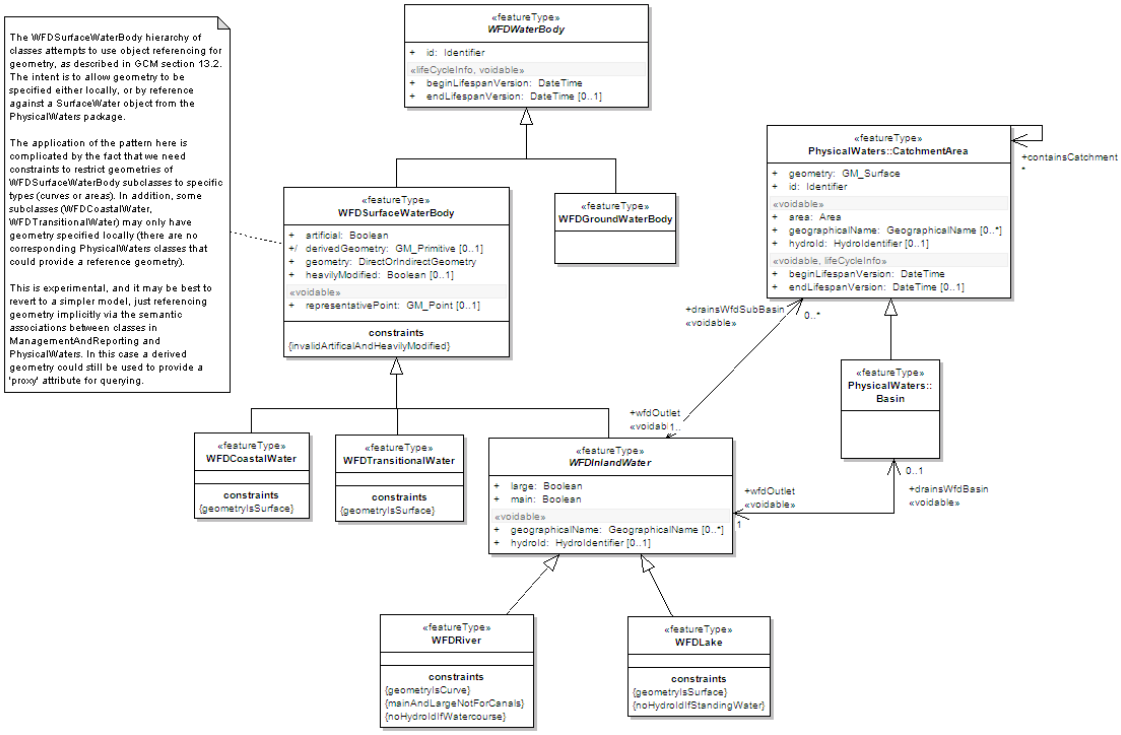

Figure 5 – Hydrography data model: ManagementAndReproting package (source: INSPIRE portal)

4. Methodology for the testing

The main part of the testing consists in performing a comparative analysis in order to identify the subset of the INSPIRE data model and feature catalogue formed by the GIS4EU datasets and the assessment of the problems that might occur at this stage of the harmonization process. The comparative analysis of the GIS4EU datasets with the INSPIRE data model comprises the following phases:

- identification of the datasets and documentation relevant to the theme hydrography.
- take as a reference the technical specifications of the INSPIRE data model for hydrography
- Detailed analysis of each dataset, comprising (1) the identification of corresponding features in the data models and (2) the identification of attributes for the features found in the previous step.
- analysis of the type of matching between features and attributes
- select the subset of the INSPIRE data model to be used in the GIS4EU project

- identify elements that do not exist in the INSPIRE data model
- identify elements not relevant in the INSPIRE data model
- report any problems found

The project deliverable D2.2 (Hobona & Jackson, 2008) that documents the data available in GIS4EU was used for the identification of the datasets relevant for the theme and for the detailed analysis of the content of each dataset. The technical specifications of the INSPIRE data model for hydrography (INSPIRE Model, 2008) were used, in the beginning using the 1st draft and, as soon as they become available (in December 2008), the consolidated UML model (2nd draft).

The comparative analysis between the project datasets and the INSPIRE data model used a matching table. The result is a list of the features and attributes corresponding in the data models being compared and the classification of that correspondence according to the following categories: A - Features/attributes from the dataset that fit on the INSPIRE data model, B - Features/attributes from dataset that are not included in the INSPIRE data model and C - Features/attributes from INSPIRE data model that are not included in the dataset. The matchings of type A are further classified into: A.1 - Direct match, A.2 - Match with some semantic or data capture differences which must be stressed and A.3 – Complex match. The features and attributes of the categories B and C are classified as relevant or not relevant for the INSPIRE directive.

The data producers performed the comparative analysis and the results of the matching process were classified according to the subtype of category A. The data producers also performed a critical analysis of the matching process where the features and attributes of categories B and C were referenced.

5. Results

The features of the INSPIRE data model that were matched with features from the GIS4EU datasets are listed according to the INSPIRE package to which they belong.

For the Network package the INSPIRE feature “WatercourseLink” is matched with features belonging to the datasets SK50, BT5M, BT50M, RLIG, RPIE and RVEN (the datasets identificators is given in the Introduction). The INSPIRE feature “WatercourseNode” is matched with features belonging to the datasets BT5M, RLIG, RPIE and RVEN.

The attributes of the feature “WatercourseLink” matched with any of the attributes of the features in the project datasets are: id, geographicalName, flowDirection, Length and centerlineGeometry. The attributes of “WatercourseNode” matched are geometry, id and hydroNodeCategory.

For the ManagementAndReporting package the only match reported by the data providers is for the RPIE dataset: “WFDSurfaceWaterBody” and the only attribute matched is id.

For the RelatedObjects package the matches found are:

- DamOrWeir: (attribute: condition), Datasets: BT5M, BT50M, EGM_PT and ERM_PT.
- SpringOrSeep: (No attributes matched), Dataset: ERM_PT.
- Embankment: (No attributes matched), Dataset: ERM_PT.
- ShorelineConstruction: (No attributes matched), Datasets: BT5M, BT50M and ERM_PT.
- AbstractPoint: (No attributes matched), Datasets: BT5M and BT50M.
- DischargePoint: (No attributes matched), Dataset: BT5M.
- Pipe: (No attributes matched), Datasets: BT5M and BT50M.
- Ford: (No attributes matched), Datasets: BT5M and BT50M.
- SubsurfaceCrossing: (No attributes matched), Dataset: BT5M.
- SurfaceCrossing: (No attributes matched), Datasets: BT5M and BT50M.
- VanishingPoint: (No attributes matched), Datasets: BT5M and BT50M.

Finally, the matchings for the features in the PhysicalWaters package are:

- StandingWater: (attributes: hydroid, geographicalName, Origin, Elevation, surfaceArea, Geometry, ID and localType), Datasets: SK50, BT5M, BT50M, RLIG, RPIE, RVEN, EGM_PT and ERM_PT.
- Watercourse: (attributes: beginLifespanVersion, endLifespanVersion, Origin, Condition, Fictitious, Level, id, geographicalName, length, LevelOfDetail, waterCourseHierarchy, localType, geometry and persistence), Datasets: BT5M, BT50M, RLIG, RPIE, RVEN, EGM_PT, ERM_PT, INSIEL and MAV.
- LandWaterBoundary: (attributes: Origin, waterLevelCategory, Geometry and id), Datasets: BT5M, BT50M, RLIG, RVEN and ERM_PT.
- Sea: (no attributes matched), Datasets: BT5M, BT50M and ERM_PT.
- Foreshore: (attribute: geographicalName), Datasets: EGM_PT and ERM_PT.
- Wetland: (no attributes matched), Datasets: BT5M, BT50M, ERM_PT.
- CatchmentArea: (attributes: area, hydroid and geographicalName), Dataset: SK50.

The INSPIRE features matched with one or more GIS4EU features are listed in table 1 that intends to summarise the results of the matching process.

Table 1 – INSPIRE features matched with GIS4EU datasets' features

Package	Matched feature types
Network	WatercourseLink, WatercourseNode
ManagementAndReporting	WFDSurfaceWaterBody
PhysicalWaters	StandingWater, Watercourse, LandWaterBoundary, Sea, Foreshore, Wetland, Riverbank, CatchmentArea
RelatedObjects	DamOrWeir, Embankment, ShorelineConstruction, DischargePoint, SpringOrSeep, Pipe, Ford, AbstractPoint, VanishingPoint, SubsurfaceCrossing, SurfaceCrossing

6. Conclusions

None of the GIS4EU datasets has a feature catalogue containing the complete INSPIRE feature catalogue on hydrography, probably because the GIS4EU data providers are mainly mapping agencies. On the other hand it was possible to find corresponding features in the INSPIRE data model for all but one of the features (Island) in the datasets ERM_PT and EGM_PT and for a large number of features in BT5M and BT50M.

The number of features directly matched in the datasets based on the Italian DbPrior10K specifications is small because these datasets only have a limited number of hydrographic features. The number of matchings increases for general purpose datasets, like those from ICC and IGP, where there is a significant number of matchings for features from the RelatedObjects package.

There are no requirements (minimum quality, mandatory features and attributes level of detail) to accomplish to be an INSPIRE dataset.

In the critical analysis performed for the matching of each dataset some data providers referred they would like the INSPIRE data model to have some specific features in order to fulfil their needs but it might be noted that the purpose of INSPIRE is to be a general purpose data model and thus, specificities might not be accommodated by it.

As a general remark about the INSPIRE data model it can be noted that there are no recommendations to solve cross-border and cross-scale issues. Moreover, the INSPIRE data model has no explicit scale associated but the outcome of the mapping between specific data models and it will inevitably reflect the inherent scales involved.

References

Hobona, G. and Jackson, M. (editors) (2008) "Data Descriptions", GIS4EU Deliverable D-2.2, Version 6.6, 20 of May 2008.

The documents referred in this paper related with the INSPIRE directive, namely, DS-D2.3, DS-D2.5, DS-D2.6, DS-D2.7 and the INSPIRE Consolidated UML Model are available, together with the text of the directive, in the document archive of the site http://inspire.jrc.ec.europa.eu/