IDM tools

NAVIGATING THE INSPIRE DATA MODEL AND GENERATING AN INSPIRE SQL DATABASE AND WFS CONFIGURATION

Alberto Belussi, Mauro Negri, Giuseppe Pelagatti

Università di Verona, Politecnico di Milano
Italy

alberto.belussi@univr.it

18 June, 2014
Outline

• Reference context and motivations
  – UML data models are the reference specification
• General architecture of the Inspire Data Model tools
  – IDM Browser: accessing to the data model by non-UML expert.
  – IDM SQL Mapper: generating INSPIRE DATABASE
  – IDM WFS Configurator: enabling the GML streams
• SQL and WFS Mapping Rules
  – Data quality preservation
  – GML encoding
• Conclusion and future work
Starting point

INSPIRE specifications have to be implemented by member states within some strict deadlines. In many countries spatial data are available as spatial databases (for example, this is the current situation in many Italian Regions). We have to start from a spatial database.

Moreover, the solution supported by INSPIRE for providing data is based on Web Feature Services that produce GML streams.
WFS are created to be queried
Reference context

Spatial databases → WFS → GML → Client Side

Query

© 2014 - Alberto Belussi
The architecture of the IDM Tools is characterized by 2 main aspects:

- the support to the frontend service is provided by an SQL database system
- the structure of the database is produced starting from the UML models (not from the XSDs of the service)
Motivations (I)

Performance

• a WFS service must be supported by a full-fledged spatial DBMS in order to obtain acceptable performance, especially in presence of one or more of the following requirements:
  – the WFS service has to process queries (possibly many queries in parallel), not only download predefined datasets.
  – the amount of data to access is large.
  – the data have to be updated through a transactional WFS.
Motivations (II)

Stability of the Data Models

the UML model will be more stable than the XML schemas (XSD files), which have been derived from it, since:

- the legally approved INSPIRE Data Specifications contain the UML model, not the XSD files.
- The XSDs files published by INSPIRE constitute only one possible encoding; other encodings may emerge in the future.
Motivations (III)

The major investment will be the *(semantic) transformation* from Production Data Models to the INSPIRE Data Model. The INSPIRE Database is derived from the UML model, thus using the INSPIRE database as the target of these transformations will safeguard these investments from changes in the encodings.

In fact, in this approach different encodings can be produced without affecting the *(semantic) transformations*.
Architecture of IDM Tools

Current IDM Tools Architecture

INSPIRE DATA MODEL TOOLS

Model Browser

SQL Mapper

WFS Configurator

Annex 1

SQL scripts

WFS configuration file

INSPIRE database (PostGIS)

WFS server (deegree)

GML stream
Model Browser

It provides a way of exploring the Conceptual Data Model in a hypertext style (or flatten mode) that is different and complementary to reading UML diagrams; this is useful not only for non UML experts, but also for experts who need to retrieve rapidly all information related to a given class.
SQL mapper
It allows one to select a set of Feature Types from the Conceptual Model and generates an SQL Script for creating a PostGis Database – called INSPIRE DATABASE - representing the selected Feature Types.
IDM tools – WFS configurator

**WFS Configurator**

It produces a Configuration File for **building a Feature Store on the INSPIRE DATABASE** enabling a Deegree WFS to access the Data contained in it and making them available through GML streams to the client applications.
Preloaded Data Model

In the current release this preloaded Data Model has been derived from the official Inspire 2010-04-26_r937.eap distribution.

The few transformations which have been applied to the INSPIRE Themes are due to the correction of some errors or to some requirements for the encoding in particular for ISO classes. The ISO package has also been simplified and renamed as “selected Types ISO TC211”.

A detailed description of these modifications is presented in the Preloaded Data Model panel of the tool.

The derived model in eap format can be downloaded from the spatialdbgroup web site (http://spatialdbgroup.polimi.it).
INSPIRE DATABASE

Goals of Inspire DB

– Supporting the implementation of WFS for spatial data having complex structure (this is the case of INSPIRE datasets)
– Guaranteeing the satisfaction of integrity constraints by database content thus ensuring high quality of provided data
– Possibly improving the performance of spatial queries and Xlink navigation
– Possibly providing the ability of managing incoming data through WFS-T service
– Being an SQL Reference Implementation
Features of Inspire Database

- **Primary key and unique constraints** derived from Data Specification are added to the SQL schema.
- **Foreign key constraints** are also specified in order to deal with referential integrity of roles and other links among tables. Constraints involving different tables are handled.
- **Cardinality constraints** are added in order to check required attributes and roles.
- **Views** can possibly be added to obtain the extent of superclasses in hierarchies.
- **Geometric domains** are specialized and constrained in order to guarantee that the represented geometries have the correct type (POINT, LINestring, POLYGON, …) and the correct coordinate dimension.
- **Spatial indices** are defined on every geometric attribute.
- **No redundancy** is introduced for the representation of a role together with its corresponding inverse role.
Snapshots of IDM tools

Welcome to the Inspire Data Model (IDM) Tools

The IDM Tools have 2 main functions:

- to support the user in exploring the Inspire Data Model in a way that is different and complementary to reading UML diagrams; the tool which supports these functions, called SQL Mapper and WFS Configurator produce an SQL Script and a WFS Configuration File, as shown by the Figure below.

- to support the design of a spatial relational database (called here Inspire Database) capable of representing the contents defined by the Inspire Specifications and the configuration of a WFS service based on the Inspire Database.

The tools which support these functions, called SQL Mapper and WFS Configurator, produce an SQL Script and a WFS Configuration File, as shown by the Figure below.

The Inspire Database is represented in the Figure as a database in the "Preloaded Data Model item of the User Guide Menu"; it is important that you read this first.

The Model Browser enables the user to explore the Inspire Data Model in a way that is different and complementary to reading UML diagrams.

© 2014 - Alberto Belussi
Snapshots of IDM browser

- Complete packages structure
- Only application schema packages
- Flat list of feature types
Snapshots of IDM browser

class: RailwayNode

Description [•] -- Definition --
A point spatial object which represents a significant point along the railway network or defines it to describe its connectivity.

Direct Properties (blue) and Inherited Properties (Yellow)

---

© 2014 - Alberto Belussi
Snapshots of IDM browser

<table>
<thead>
<tr>
<th>Direct Properties (blue) and Inherited Properties (Yellow)</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;voidable&gt;formOfNode(codeList: FormOfRailwayNodeValue)[1..1]</td>
</tr>
<tr>
<td>&lt;voidable&gt;beginLifespanVersion(Date Time)[1..1] FROM: NetworkElement [Network::NetworkElement]</td>
</tr>
<tr>
<td>&lt;voidable&gt;endLifespanVersion(Date Time)[0..1] FROM: NetworkElement [Network::NetworkElement]</td>
</tr>
<tr>
<td>&lt;voidable&gt;validFrom(Date Time)[1..1] FROM: TransportNode [Common Transport Elements::TransportNode]</td>
</tr>
<tr>
<td>&lt;voidable&gt;validTo(Date Time)[0..1] FROM: TransportNode [Common Transport Elements::TransportNode]</td>
</tr>
<tr>
<td>&lt;voidable&gt;geographicalName(datatype: GeographicalName)[0..1] FROM: TransportObject [Common Transport Elements::TransportObject]</td>
</tr>
<tr>
<td>&lt;voidable&gt;inspireId(datatype: Identifier)[0..1] FROM: NetworkElement [Network::NetworkElement]</td>
</tr>
<tr>
<td>geometry(GM_Point)[1..1] FROM Node [Network::Node]</td>
</tr>
<tr>
<td>R &lt;voidable&gt;speakStart[0..*]Link FROM Node[Network::Node] inverse startNode[0..1]</td>
</tr>
<tr>
<td>R &lt;voidable&gt;speakEnd[0..*]Link FROM Node[Network::Node] inverse endNode[0..1]</td>
</tr>
<tr>
<td>R &lt;voidable&gt;inNetwork[1..<em>]Network FROM NetworkElement[Network::NetworkElement] inverse elements[0..</em>]</td>
</tr>
</tbody>
</table>
Property Flattening in IDM browser

Attribute with structure (Datatype)

Role hidden in a Datatype

Datatype hierarchies produce different alternative possible extensions of the attribute structure

List of all roles including the hidden ones

© 2014 - Alberto Belussi
Future Work

If the approach adopted by the IDM Tools will encounter the interest of stakeholders, we plan to develop it further in several directions:

• **Validation of the INSPIRE Data Model**: the procedures which have been applied for preloading the UML model can be transformed in a distributable tool for UML Schema Validation.

• **Support for the transformation of data from existing datasets to the INSPIRE Database**: some work in this direction, considering as source database the Italian “National Core” - the Italian standard for GeoTopographic Databases - , is underway in cooperation with CISIS (a consortium of Italian Regions).
Download it now!

You can download the IDM Tools software at the following URL:

http://spatialdbgroup.polimi.it

and build your INSPIRE DATABASE

THANK YOU

QUESTIONS?
Additional slides

In the following slides some additional issues are illustrated.
Example: a DB generated by deegree

- Primary key completely useless
- no Foreign key constraint
- no Foreign key constraint and redundancy
- no Foreign key constraint and redundancy
1) General Architecture

- **Source Databases**
  - SQL scripts
  - INSPIRE DATABASE (e.g. PostGis)

- **SQL mapping generator**
  - Schema in EAP format
  - INSPIRE SCHEMA LOADER

- **INSPIRE DATA MODEL Tools** (Inspire Data Specifications)
- **WFS Configuration**
- **VALIDATOR**
- **GML datasets**
- **WFS**

© 2014 - Alberto Belussi
IDM tools and INSPIRE MIF

- Updates schema in EAP format
- Errors (violations of IDM rules)
- Statistics and Analysis by different navigation
- INSPIRE SCHEMA LOADER
- INSPIRE DATA MODEL Tools (Inspire Data Specifications)
Comparison with database generated from XSD files (Deegree)

- Primary key and unique constraints cannot be always derived from XSD syntax.
- Foreign keys cannot be specified in XSD syntax, thus are not defined in SQL schema derived from XSD.
- Cardinality constraints and domain constraints can be derived from XSD syntax, but are not always specified in SQL schema.
- Geometric types cannot always be derived from XSD syntax.
- Spatial indices are a possible choice when a geometric attribute is defined, however schema generators do not always produce them.
- Redundancy in role/inverse role implementation cannot be avoided starting from XSD syntax, since inverse roles are not specified.