Abstract

Many Legally Mandated Organisations (LMOs) and other data suppliers who are providing data within the INSPIRE themes will retain their own schemas for the data that they collect, manage and publish. For INSPIRE compliance however, they must also ensure that there is a Transformation Service available to transform the data to the appropriate INSPIRE schema. This service may be provided by the supplier; by the national government concerned; or by a third party. It may be an offline service with the INSPIRE compliant datasets then published on the web for the INSPIRE GeoPortal (and others) to access; or it may be available as a web service - preferably capable of being invoked automatically at the time that data is requested. This paper describes the State of the Art Analysis and the subsequent development of Technical Guidance for INSPIRE Transformation Network Services, as carried out by a consortium of RSW Geomatics, 1Spatial, and Rob Walker Consultancy for the EC Joint Research Centre. It will be of use to any organisation implementing a transformation from one schema to another and specifically to the GML schemas required by the INSPIRE Implementing Rules. This paper concludes that

* This work is licensed under the Creative Commons Attribution-Non commercial Works 3.0 License. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-nd/3.0/ or send a letter to Creative Commons, 543 Howard Street, 5th Floor, San Francisco, California, 94105, USA.

DOI: 10.2902/1725-0463.2011.06.art1
INSPIRE Schema Transformation Network Services are feasible with adherence to international standards, some already specified in INSPIRE regulations, together with the use of discreet standards based components that can be supplied by a variety of vendors. The Technical Guidance (TG) should help to move INSPIRE into the mainstream of current developments on the World Wide Web and spread the interoperability of geo-spatial data to a wider audience.

**Keywords:** INSPIRE, Schemas, Transformation, Guidelines

1. **INTRODUCTION**

Given the number of data providers involved and the variety of schemas in which the relevant datasets are held, it is important that some technical guidance is available from the INSPIRE programme about meeting this requirement. JRC contracted RSW Geomatics, 1Spatial and Rob Walker Consultancy to prepare Technical Guidance for completion in July 2010.

The first part of the contract was a State of the Art Analysis (SAA) (Beare et al, 2010), to evaluate:

- **Schema Description Languages** compatible with the INSPIRE data specification, source schema and mapping languages.
- **Model Mapping Languages** that define, represent and handle mappings between source and target schema, including inherent limitations.
- **Existing Transformation Tools** that support schema transformation as well as more straightforward coordinate and format transformations.
- **Enterprise Architectures** that provide for deployment of a publicly available transformation service.

The SAA report provides a summary of the key findings of the analysis and is the basis for the production of the Technical Guidance (Howard et al, 2010) which covers:

- **Architectural Goals & Constraints** including the Implementing Rules and the INSPIRE principles as well as the current information systems environment and in particular the availability of internationally recognised standards.
- **Use Case View** interpreted from the INSPIRE Regulations (European Commission, 2009).
- **Logical View** - an overview of the logical components involved including reference to the standards currently favoured for maximum flexibility.

---

• **Data View** with specific recommendations on schema description and model mapping languages.

• **System Qualities** including vendor neutrality, error handling, and interoperability.

The contract also required the demonstration of a prototype with actual data from different countries and from at least three of the INSPIRE themes and the production of a multi-lingual video to explain the Technical Guidance to those involved. The prototype was demonstrated at the INSPIRE conference in Krakow on 22 June 2010\(^2\) and the components are briefly outlined in this paper; the video is now available on the INSPIRE web site\(^3\).

2. **ARCHITECTURE & DATA CONCEPTS**

INSPIRE aims to provide integrated geospatial information services which enable user to identify and access that information from a wide range of sources for a variety of applications. All data and services will be described by metadata enabling humans and software applications to discover specific relevant datasets and service instances in the infrastructure and to decide whether they are fit for any particular application. For context, the INSPIRE conceptual technical architecture (European Commission, 2007) is reproduced in Figure 1.

![Figure 1: INSPIRE Technical Architecture Overview](http://inspire.jrc.ec.europa.eu/index.cfm/pageid/5)

---


\(^3\) [http://inspire.jrc.ec.europa.eu/index.cfm/pageid/5](http://inspire.jrc.ec.europa.eu/index.cfm/pageid/5)
Guidelines for INSPIRE data specifications\(^4\) provide detailed technical provisions for data interoperability, but it is unrealistic to expect spatial data providers to migrate all of their local operating environments to be INSPIRE compliant – at least in the short term. Therefore, to achieve compliance, transformation services may be established – by data providers or others – to transform data from local schemas into INSPIRE schema(s). It will be for the data provider to decide whether this transformation should be an in-house process, prior to making the data available, or whether the original data should be provided accompanied by a dynamic, open access transformation service.

The Technical Guidance being offered through this project is primarily aimed at the open access service, but elements of the guidance may also be applicable to others. Typical workflows for using INSPIRE to access disparate data sources are illustrated in Figure 2. Schema transformation services are required ‘up-front’ and will most likely be integrated with relevant co-ordinate transformations.

*Figure 2: Illustration of a Possible Transformation Process Sequence*

- Transformation is the key to INSPIRE interoperability and effective data re-use.
- Schema transformation provides the foundation for other processes to operate in a common framework.

\(^4\) http://inspire.jrc.ec.europa.eu/index.cfm/pageid/2
The scope of the Technical Guidance for INSPIRE Transformation Services is limited to schema transformation, and is illustrated in Figure 3. The ‘Transformation Web Service’ represents the generic web service, for which each type of transformation could be implemented – this paper deals specifically with schema transformations.

**Figure 3: Conceptual Scope of Technical Guidance for INSPIRE Transformation Service**

The Draft Implementing Rules for Transformation Services for INSPIRE (European Commission, 2009) define a set of service operations and accompanying parameters for the Transformation Web Service which form the ‘Web Service Interface’ between the web service and its ‘Transformation Service Consumer’ (a client application or other service component).

The Technical Guidance considers:

- Schema Descriptions.
- Model Mappings between source and target schemas.
• Schema Transformation in a service oriented architecture environment.
• Paradigms existing within the geospatial sector and in general IT.
• Parameters for service operations – in-line (byValue) or with identifier/pointer (byReference).

3. STATE OF THE ART ANALYSIS

3.1. Schema Description Languages

Transformation services require definitions of the source and target schemas in order to be able to apply mappings to the actual datasets. The candidate languages, based on previous experience and user feedback, are identified in Table 1.

<table>
<thead>
<tr>
<th>Name / Version of Language</th>
<th>Originator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unified Modelling Language (UML) 2.2⁵ / XML Metadata Interchange (XMI) 1.1⁶</td>
<td>OMG</td>
</tr>
<tr>
<td>XML Schema Definition (XSD) 1.1⁷ / Geography Markup Language (GML) 3.2.1⁸</td>
<td>W3C</td>
</tr>
<tr>
<td>Resource Description Framework (RDF) 1.0⁹ / Web Ontology Language (OWL) 2.0¹⁰</td>
<td>W3C</td>
</tr>
</tbody>
</table>

Each of these languages was evaluated as strong, acceptable or weak under the following headings:

• **Expressiveness** – ability to represent all concepts required for schema definition
• **Mapping Compatibility** – suitability for use with the mapping language through a well-defined process.
• **Web Compatibility** – proven capability for web service applications
• **Tool Support** – for editing schema descriptions and for inferring initial schema descriptions from existing data storage.

⁵ http://www.omg.org/spec/UML/2.1.2/
⁶ http://www.omg.org/spec/XMI/2.1.1/
⁷ http://www.w3.org/TR/xmlschema-0/
⁸ http://www.opengeospatial.org/standards/gml
⁹ http://www.w3.org/TR/rdf-schema/
¹⁰ http://www.w3.org/TR/owl-guide/
• **Technology Independence** – neutrality of vendors and data-encoding formats.
• **Intuitiveness** – simple and concise representation for data modellers.

Table 2 Strengths and Weaknesses of Schema Description Languages

<table>
<thead>
<tr>
<th>Criteria</th>
<th>UML/XMI</th>
<th>XSD/GML</th>
<th>RDF/OWL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expressiveness</td>
<td>Strong</td>
<td>Acceptable</td>
<td>Acceptable</td>
</tr>
<tr>
<td>Web Compatibility</td>
<td>Strong</td>
<td>Strong</td>
<td>Strong</td>
</tr>
<tr>
<td>Tool Support</td>
<td>Strong</td>
<td>Weak</td>
<td>Acceptable</td>
</tr>
<tr>
<td>Technology Independence</td>
<td>Strong</td>
<td>Weak</td>
<td>Strong</td>
</tr>
<tr>
<td>Intuitiveness</td>
<td>Strong</td>
<td>Weak</td>
<td>Weak</td>
</tr>
</tbody>
</table>

Schema description languages and model mapping languages need to be evaluated together. UML/XMI is a strong candidate, but proved to be problematic in practice, because there is no consistent format for export of XMI documents from UML models. XSD/GML is generally considered a logical format, but its physical structure is verbose and required extra effort when developing mappings against it. RDF/OWL is an ontology language designed for the Semantic Web.

Subsequently, in the Technical Guidance, XSD/GML is preferred. The rationale for this was based on the key requirement for a schema description language to be interoperable with the mapping interchange format (which, as discussed in section 3.2 is the Rules Interchange Format). As the schema description language is an interchange format, it is important that it is both sufficient for the service and practically usable by service clients, typically by conversion from an existing schema description. XML Schema is most closely aligned to the needs of the Schema Transformation Network Service, since it is designed for the strict definition of concrete datatypes. It has good support for common data modelling techniques such as structured types containing inner-typed properties and cardinalities, which are not naturally part of RDF Schema or OWL, although they can be emulated. UML can be used to express a range of modelling abstractions (including physical models) but it cannot be used to express the actual data content.

3.2. **Model Mapping Languages**

There is no standard meta-language for model mappings and few standards exist for schema transformations and schema mapping encodings. This makes it difficult to implement any schema transformation service in a platform-neutral manner but a suitable model mapping language must be selected or designed for
use in the transformation service request parameters. Several ISO, W3C and OGC ‘de jure’ standards were evaluated. Those chosen are based on experience and on user feedback.

Table 3: Candidate Model Mapping Languages

<table>
<thead>
<tr>
<th>Language</th>
<th>Version or Date</th>
<th>Originator</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extensible Stylesheet Language for Transformations (XSLT)(^{11})</td>
<td>2.0</td>
<td>W3C</td>
<td>Standard</td>
</tr>
<tr>
<td>Web Ontology Language (OWL)(^{12})</td>
<td>2.0</td>
<td>W3C</td>
<td>Standard</td>
</tr>
<tr>
<td>Rule Interchange Format (RIF)(^{13})</td>
<td>1.0</td>
<td>W3C</td>
<td>Standard</td>
</tr>
<tr>
<td>Semantic Web Rule Language (SWRL)(^{14})</td>
<td>21/05/2004</td>
<td>W3C</td>
<td>Standard</td>
</tr>
<tr>
<td>Query/View/Transform (QVT)(^{15})</td>
<td>1.0</td>
<td>OMG</td>
<td>Standard</td>
</tr>
<tr>
<td>Common Logic (CL)(^{16})</td>
<td>ISO/IEC IS 24707:2007</td>
<td>ISO</td>
<td>Standard</td>
</tr>
<tr>
<td>Ontology Mapping Language (OML)(^{17})</td>
<td>06/10/2005</td>
<td>DERI OMWG</td>
<td>Specification</td>
</tr>
<tr>
<td>Rewerse II Rule Markup Language (R2ML)(^{18})</td>
<td>0.5</td>
<td>WGI1</td>
<td>Specification</td>
</tr>
<tr>
<td>Tefkat(^{19})</td>
<td>2.1.0.lawley 266</td>
<td>DSTC Australia</td>
<td>Other</td>
</tr>
</tbody>
</table>

RIF, QVT and XSLT were selected for more detailed evaluation.

RIF is about to be adopted as a W3C standard, has a formal theoretical basis, and is aligned with OWL and RDF which are becoming central to the Semantic

\(^{11}\) http://www.w3.org/TR/xslt20/
\(^{12}\) http://www.w3.org/TR/owl-guide/
\(^{13}\) http://w3c.org/TR/rif-overview/
\(^{14}\) http://www.w3c.org/Submission/SWRL/
\(^{15}\) http://www.omg.org/spec/QVT/1.0/
\(^{16}\) http://common-logic.org
\(^{17}\) http://www.omwg.org/TR/d7/d7.2/
\(^{18}\) http://rewerse.net/l1/oxygen.informatik.tu-cottbus.de/rewerse11/@q=node_2f6.htm
Web and open data initiatives. There are a number of experimental implementations, including products from Oracle and IBM.

QVT is based on the MOF/UML family of languages arising from years of practical Model Driven Architecture experience in real-world implementations in general information modelling domains. It has tool support from the Eclipse Modelling Framework e.g. M2M, IBM tools.

XSLT is limited to XML processing but is familiar and widely adopted over a decade, with two versions of specification released. Its strengths and weaknesses are well known and it has relatively straightforward extensibility via Java or other 3/4GL languages.

All of these languages were deemed fully acceptable for Expressiveness but showed distinct differences for implementation criteria.

**Table 4: Comparison of Model Mapping languages for Implementation Criteria**

<table>
<thead>
<tr>
<th>Implementation Criteria</th>
<th>RIF</th>
<th>QVT</th>
<th>XSLT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Technology independence</td>
<td>Strong</td>
<td>Strong</td>
<td>Strong</td>
</tr>
<tr>
<td>Practical feasibility</td>
<td>Acceptable</td>
<td>Strong</td>
<td>Weak</td>
</tr>
<tr>
<td>Intuitiveness</td>
<td>Acceptable</td>
<td>Strong</td>
<td>Weak</td>
</tr>
<tr>
<td>Manageability</td>
<td>Strong</td>
<td>Acceptable</td>
<td>Acceptable</td>
</tr>
<tr>
<td>Web compatibility</td>
<td>Strong</td>
<td>Strong</td>
<td>Strong</td>
</tr>
<tr>
<td>Logical portability</td>
<td>Strong</td>
<td>Acceptable</td>
<td>Weak</td>
</tr>
<tr>
<td>Custom Extensions</td>
<td>Strong</td>
<td>Strong</td>
<td>Acceptable</td>
</tr>
</tbody>
</table>

The analysis therefore suggests that XSLT is a much weaker contender than either RIF or QVT.

In the Technical Guidance RIF was subsequently preferred. RIF has been designed primarily to be a rule/action interchange language, rather than an implementation language. It is mathematically rigorous, and has been disseminated by a well-respected international standards body (W3C). RIF has been designed by a group of experts representing the distilled experience of the Business Rules Management Systems industry. It is appropriate to the challenge
of schema transformation, and indeed it is one of the use cases for which the format was designed.

3.3. Existing Transformation Tools

Investigation of the existing tools helped to validate the practicality of requirements defined in the Implementing Rules (IR) and also to identify possible constraints for the Technical Guidance, to ensure that the transformation services can be implemented by multiple vendors.

The following vendors/distributors responded to a questionnaire as part of the State of the Art Analysis.

**Table 5: Vendors/distributors of transformation tools**

<table>
<thead>
<tr>
<th>Vendor / distributor</th>
<th>Commercial or R &amp; D</th>
<th>Name of tool</th>
<th>Tool Version</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAFE Software</td>
<td>Comm.</td>
<td>FME Server</td>
<td>2010</td>
</tr>
<tr>
<td>Snowflake Software</td>
<td>Comm.</td>
<td>GO Publisher</td>
<td>1.4</td>
</tr>
<tr>
<td>interactive instruments GmbH</td>
<td>Comm.</td>
<td>XtraServer</td>
<td>3.2</td>
</tr>
<tr>
<td>1Spatial</td>
<td>Comm.</td>
<td>Radius Studio</td>
<td>2.1.0.15</td>
</tr>
<tr>
<td>lat/lon GmbH</td>
<td>R&amp;D</td>
<td>Deegree WPS</td>
<td>3.0</td>
</tr>
<tr>
<td>Talend</td>
<td>Comm.</td>
<td>Talend Integration Suite</td>
<td>Not specified</td>
</tr>
<tr>
<td>Humboldt</td>
<td>R &amp; D</td>
<td>Humboldt Alignment Editor/Conceptual Schema Translation Service</td>
<td>HALE 2.0.0-M1, CST 1.0.0-RC1</td>
</tr>
<tr>
<td>AuScope</td>
<td>R &amp; D</td>
<td>AuScope Grid (uses GeoServer)**</td>
<td>Not specified</td>
</tr>
</tbody>
</table>

An online survey obtained information for each product about its capabilities for schema transformation at different levels. The levels were introduced to describe different types of functionality that will be required in order to transform schema of varying complexity. When the source schema is closely aligned to the target schema, a lower level of transformation functionality will be required. Each level incorporated all functionality from earlier levels, i.e. if a transformation service supports functionality in level n, it should also support all functionality in level n–1. Further information can be found in the SAA, but in brief, the levels were:
• Level 1 - Renaming classes and attributes.
• Level 2 - Simple attribute derivation.
• Level 3 - Aggregating input records.
• Level 4 - Complex derivation and dynamic type selection.
• Level 5 - Deriving values based on multiple features.
• Level 6 - Conflation and model generalisation.

Subsequently, further information was obtained from public websites and in some cases by correspondence.

This survey demonstrated that there are no widely used standards for schema descriptions or model mappings, highlighting the importance of this project. Most respondents use some form of proprietary language for schema descriptions and model mappings.

Most respondents claimed support for all levels of transformation functionality. This implies that transformation services for INSPIRE should be feasible, with a wide choice of potential tool vendors. It also enables the Technical Guidance to select a model mapping language that is sufficiently expressive to describe all mappings that are likely to be required. Most tools support GML, Oracle Spatial and ESRI Shape files and have optional support for syntactic validation of the model mapping, to ensure that it meets the rules for the target schema. However, few support semantic validation to ensure that the data is fully compliant with the data specifications.

Most tools can run in several environments, including combinations of Linux, Windows, 32bit and 64bit. Desktop application and batch processes are frequently provided, with some tools providing web accessible interfaces. The majority of tools claim support for scalable processing, including processing of multiple simultaneous requests – so it is likely that the performance requirements defined in the IR will be met by many tools.

4. CANDIDATE ARCHITECTURES

Any transformation network service must be considered in relation to the overall INSPIRE architecture to determine any impact on the request and response parameters or functionality of the service.

The INSPIRE Directive requires that transformation services should be “available to the public and accessible via the Internet or any other appropriate means of telecommunication”. The Draft IR (European Commission, 2009) proposes several potential service architectures, each of which has various strengths and weaknesses, depending on the technical and business processes in which they will operate. The application may need to interact directly with the Download
These service architectures have been evaluated against the following quality of service criteria:

- Performance
- Scalability
- Reliability, resilience and availability
- Flexibility
- Extension of functionality
- Cost

Additional factors that may need to be considered will depend on the architecture chosen for deployment and are outside the scope of this project. These will include: platform infrastructure and deployment; service monitoring; maintenance; installation; testing; upgrade; localisation; network traffic/bandwidth; service level agreement and contractual management; and security.

4.1. Independent Service Node

Illustrated in Figure 4, an application calls the Transformation Network Service (TNS) to transform the query, a Download Service to obtain the data, and finally the TNS again, so that the result set can be transformed into the standard format. This approach scores well for flexibility but is weak for all other criteria.
4.2. Tightly coupled proxy facade

Illustrated in Figure 5, a client calls the TNS and there is no direct access to the content access service. The TNS therefore orchestrates any combination of calls to relevant content access services (e.g., WFS and WMS) to fulfil the request. This approach was considered strong in terms of flexibility, scalability and extensibility. However, the approach is weak for reliability, performance and opportunities for quality assurance.
4.3. Encapsulation in Other Services

Illustrated in Figure 6, a client will ask for data from a content access service (e.g. download service), specifying the input and output format and mapping required. The service will compile the dataset and then invoke the TNS on the way back, so that the content received by the client is in the standard format. In this use case, the actors are the target content access service and the TNS. There is no direct call to the TNS by the client. Strengths and weaknesses are similar to those for the tightly coupled proxy facade. However, the download service must be implemented / modified to have knowledge of the transformation service. In many cases, a download service might not need to use a transformation service, so this is adding additional complexity and reducing the possibility for re-use of standard components.
4.4. Bulk Transformation & Caching

There is alternative approach to transformation services not mentioned in the IR. Transformation of source data to the INSPIRE schema is pre-computed and the transformed data resulting from this is then cached. The download service and client applications communicate with this data cache in the INSPIRE model and do not need to know about the existence of data in any other format. However, the bulk transformation service is considered as an online service in which the first requester will incur the penalty of the additional time required to wait for the results of the transformation, and subsequent requesters will received the pre-transformed dataset.

Strengths of this approach are in performance, scalability, flexibility and extensibility. Weaknesses could be lack of currency for volatile data and the change in architecture required to encompass an additional dataset.
5. ARCHITECTURAL GOALS AND CONSTRAINTS

The following generic issues - high level features, technical risks or overarching constraints - for Schema Transformation Network Services are such that they will have a significant architectural impact.

5.1. EC Regulations

A ‘Network Service of type Transformation’ must meet the provisions of the INSPIRE Regulation and in particular needs relating to logical schemas.

5.2. Mapping Flexibility

The TG aims to define an interface that is rich enough to allow implementations supporting transformations from a wide variety of source schemas into equivalent INSPIRE schemas. However, individual implementations will have limitations imposed by supporting technologies and may only support a subset of the source and INSPIRE schemas.

The INSPIRE schemas are defined by the INSPIRE Regulations as guided by the INSPIRE data specifications. They should include all themes (including those that are currently, or will in future be, under development). The source schema is determined by data providers following their standard data capture and schema development processes. It is therefore desirable to configure Schema TNS to work with a wide variety of source and target schemas.

5.3. Open Interfaces

In order to enable interoperability within INSPIRE based projects, Schema TNS should be based on the common interface specification defined in the INSPIRE Regulations including metadata requirements. These must not be tied to particular software – commercial or open-source. It should be possible to create multiple implementations of Schema TNS, each using a different underlying transformation engine.

If the transformation engine has to be replaced (for example, due to performance, cost or other features) or, alternatively, another Schema TNS is to be consumed, then it should be possible to do so without re-writing the schema mapping definitions.

5.4. Statelessness

There should be no need for the client to perform any other interaction with the service – the service should be stateless. All information required to perform a transformation should be provided in the initial operation request. The service
does not therefore store the transformed data but transfers it to a location (either WFS-T or FTP site) nominated by the client in the initial operation request.

This stateless design makes the service more amenable to load balancing (a process whereby requests to a single virtual service are routed automatically to one of several actual services) which improves scalability and resilience. It also simplifies data management and minimises licensing issues associated with the storage and processing of the spatial data.

5.5. Separation of Control Messages from Data Transfer

XML encodings of web service requests and responses are typically small (a few kilobytes or less) and modern web service development platforms have been designed on this assumption and process requests in memory. However, spatial datasets are very large – especially when encoded in XML (many hundreds of megabytes). If the spatial datasets were embedded directly in either the request or the response messages of a Schema TNS, they would be too large to be processed directly in memory.

The Technical Guidance therefore recommends that spatial datasets are never passed directly through the service interface but are, instead, passed by reference. This enables the most appropriate technology to be used when handling the actual transfer of the spatial data.

Another advantage of this recommendation is that a complete record of the messages sent to and received from a service can easily be recorded for audit or debugging. This would be very difficult to achieve if spatial datasets were passed by value through the interface. The same advantage applies to spatial dataset schemas when expressed as GML application schemas and also mapping definitions.

5.6. Schema-Agnosticism

The TG describes an interface that is entirely schema-agnostic – it embeds no knowledge of the structure of any source schema or any INSPIRE schema in the request message. This enables the interface definition to remain constant even when these schemas change. Therefore, all data transformed by the system is considered equal; the system will handle identifiers, data and metadata in the same manner – all INSPIRE identifiers, data and metadata must be derived from the identifiers, data and metadata available in the source datasets.

5.7. Automated Process

Schema TNS are designed to be fully automated within an integrated operating environment, including features such as orchestration of services, security, rights management and quality of service provisions.
6. PROTOTYPE

To demonstrate the practical feasibility of the Technical Guidance, a prototype demonstrator was developed\(^\text{20}\). The prototype is based exclusively on the Technical Guidance, i.e. no assumptions or service parameters have been added. The Technical Guidance is expected to provide the authoritative source for the operation signatures, parameter data types and modes of transport of parameter data, without requiring reference to any other document. The objective of the prototyping phase was to prove the principles of the Technical Guidance.

To deploy the components needed to demonstrate the prototype Transformation Network Service a test environment was configured to include pre-existing software tools from four separate organisations and open source communities:

- GeoServer\(^\text{21}\), an open source Web Feature Service (WFS), for publishing source data/schema;
- Humboldt Alignment Editor (HALE)\(^\text{22}\), an open source tool for defining schema mappings;
- Radius Studio™\(^\text{23}\), a commercial geospatial rules engine for performing the transformations;
- TatukGIS Viewer\(^\text{24}\), a free to use application for visualising transformed data.

It is possible, by exercising the primary recommendations made by the Technical Guidance, to bring these independent and disparate tools together to provide a flexible and effective transformation service. This involves use of international standards for the exchange of data, logical schema descriptions and schema mappings. Of particular note are the recommendations to use:

- Geography Markup Language (GML)\(^\text{25}\) standard, from Open Geospatial Consortium (and already endorsed by INSPIRE), for data and schema descriptions;
- Rules Interchange Format (RIF)\(^\text{26}\), a recently adopted standard from World Wide Web Consortium, for schema mapping definitions.

These components, configured around an XML repository based on open standards, are illustrated in Figure 7. Note that WFS-T is a transactional Web Feature Service.

\(^{21}\) http://geoserver.org/display/GEOS/Welcome
\(^{22}\) http://community.esdi-humboldt.eu/wiki/2
\(^{23}\) http://www.1spatial.com/software/index.php
\(^{24}\) http://www.tatukgis.com/Products/EditorViewer.aspx
\(^{25}\) http://www.opengeospatial.org/standards/gml
\(^{26}\) http://www.w3.org/2005/rules/Overview.html
Any of these components can be replaced with others capable of working with the well-defined standards being used – a true ‘plug and play’ scenario that should be both flexible and relatively future proof. To test and demonstrate the prototype service, six test datasets were kindly provided by six thematic data partners, presenting a broad coverage of transformation scenarios in respect of three INSPIRE data themes (Cadastral Parcels, Hydrography and Transport Networks). Belgian Cadastre, Dutch Kadaster, National Land Survey Sweden, Statens Kartverk Norway, Ordnance Survey Ireland, Land & Property Services Northern Ireland.

Figure 7: Prototype with Components Demonstrating the Technical Guidance

The prototype shows that by following the Technical Guidance, transformation is possible using many different tools if industry standards are followed throughout the process. Although data providers or publishers can carry out transformations themselves, they can also be offered as Network Services available to a broader community. The recommendations of the Technical Guidance are considered applicable to a variety of deployment scenarios, not just network service environments. Member States and their data providers will, of course, decide for themselves exactly where and how these transformations are delivered. Videos
explaining the TNS in English, French and German are available on the INSPIRE web site.

7. CONCLUSIONS

The contract has produced a useful State of the Art Analysis and the first version of Technical Guidance for INSPIRE Schema Transformation Network Service to assist implementation of the Network Service Implementing Rules using INSPIRE compliant Schemas from the individual Data Specifications.

As part of the State of the Art Analysis conducted earlier in the project (Beare et al, 2010), a variety of existing tools were assessed with regard to their potential to offer transformation capabilities. A common drawback is that they are not based on standards and are therefore not interoperable, encouraging vendor lock-in.

The Technical Guidance has been proven to address this problem successfully, through the adoption of open standards to manage the interoperable interchange of schema descriptions and model mapping definitions.

The Technical Guidance makes several recommendations which have been implemented in a prototype Transformation Network Service which will be promoted in short video that is under production at the time of writing.

This demonstrates that this Network Service can be achieved with:

- Vendor neutrality
- Decoupling of ‘model mapping definition’ from ‘transformation execution’
- Defined interfaces that have the best chance of meeting current and future INSPIRE schema transformation scenarios

and using industry standard languages:

- **GML/XSD** – for source & target schema descriptions.
  - Established standards with good tool support
  - Physical data model closely tied to data itself
  - Reinforcing INSPIRE use of GML
- **RIF (Rules Interchange Format)** – as output from rule authoring process
  - Rigorous academic backing
  - Final stages of W3C adoption

The prototype demonstrated that with vendor and open source community support for open standards, pre-existing tools can be adapted to be interoperable and capable within a domain previously dominated by proprietary solutions.

We anticipate that other technology providers/developers could similarly adapt their tools to support the open standards. Therefore, it would be possible to use other technologies, as required, to fulfil the roles of the components identified in this work. This would create flexibility for service developers and system integrators to use appropriate technology without requiring re-investment in describing schemas and/or model mappings. It also shows that, rather than remaining in a niche ‘geoweb’, spatial datasets are moving into the generic standards based environment of the World Wide Web.

Even with the modest infrastructure used by the prototype test, reasonable transfer rates were achieved. The stateless design of the Schema Transformation Network Service means that adding additional processing resources can smoothly scale this throughput. Radius Studio, for example, is backed onto a scalable grid-based technology, which allows extra resources to be added to the engine to increase performance. This grid-based technology also delivers fail-over support, which in turn is essential in providing highly available services. Other vendors may use grid-based or other mechanisms to deliver similar non-functional characteristics; the prototype has proved that this kind of enterprise technology is compatible with the interface defined by the Technical Guidance.

REFERENCES


– HALE-to-RIF open source plug-in available from this project, extending the work of Humboldt.
