THE EUROPEAN LOCATION FRAMEWORK (ELF), INTEROPERABILITY SOLUTIONS FOR EUROPEAN PUBLIC ADMINISTRATIONS AND THE INSPIRE DIRECTIVE.

Major challenges and lessons learnt for delivering authoritative, interoperable, cross-border geospatial reference data for a European coverage.

Nathalie Delattre, National Geographic Institute of Belgium

Scope

The European Location Framework builds a geospatial data infrastructure and provides pan-European interoperable reference data and services from national information assets enabling users to build their work on it.

ELF has set up geo-processing guidelines in order to produce geospatial data, interoperable, harmonised, cross border and of high quality. These geo- p*r*ocesses encompass the data transformation, the data aggregation and edge-matching, the data quality evaluation and conformance testing, the quality management, the data generalization, the incremental updates, and the visualisation.

The paper describes the best practices, major challenges and lessons learnt when applying the INSPIRE interoperability principles to those geo-processes for obtaining the relevant ELF data. The European Location Framework (ELF)[1] has investigated some aspects of semantic and technical interoperability.

Semantic Interoperability

The semantic interoperability covers the requirements in data interoperability.

One objective of the ELF project is to ensure that the national datasets, available and accessible on the ELF platform will comply with INSPIRE and moreover provide cross-border consistency and a ‘certain’ degree of data interoperability that will allow their use in a broad European perspective. It is therefore important to determine which requirements in the matter of data interoperabilitywhich must be reached to allow their specific use in a pan-European framework.

ELF has conceived several degrees in data interoperability and has set up procedures based on best practices examples to achieve them. In a way, we could define a graduated scale evaluating the degree of interoperability of spatial datasets.

[1] European Location Framework (ELF):Launched in March 2013, the project will run for three years and deliver a pan European cloud platform and web services to build on the existing work of the INSPIRE Directive and enable access to harmonised data in cross border application. http://www.elfproject.eu/

Figure 1: Degrees of interoperability of the ELF spatial datasets

This graduated scale shows the degrees of interoperability in such a way that the previous degree must be achieved in order to be able to reach the following one.

Basically, the degree of interoperability "INSPIRE compliance" sets up the technical framework, laying the foundations for interoperability, in a service-based architectureand provide the European frame for sharing interoperable data across domains through services. INSPIRE sets standards and obligations, to share and provide existing data but it does not create an obligation to update existing data nor to produce new data.Data gaps and out-datedness will not be solved by INSPIRE.

The following step to progress in the data interoperability is to define a core data content at European level. The 'core' part of a data is defined by the minimum data content required by most or at least by basic applications, and user requirements, and whose data collection and maintenance is considered as feasible and achievable at reasonable costs. Typically, this might be the mandatory feature types and attributes that will be produced and made available in all countries in order to avoid gaps in the European coverage.

Core data content may be defined independently from any Level of Detail (LoD) implying that, ideally, the data should be collected once at the most detailed level.

Theestablishment of a core data content and the determination of what should be mandatory and available in a European coverage is not a one-shot concept but is usually driven in a step by step approach based on compromise on what the data producers can collect and maintain at affordable costs and meet the main identified users' requirements and applications.

A core data content may be portrayed or represented at different LoDs and therefore be available by means of several spatial datasets.

ELF has classified the spatial datasets according to several LoDs:

LoD	Scale range	Thematic scope
Master Level 0	Largerthan 5k	Cadastral Parcels, Buildings, Addresses
Master Level 1	5k – < 25k	ELF Topo (Admin Units, Hydro,

		Transport, Elevation, GeoNames, etc.)
Master Level 2	25k – <100k	ELF Topo generalised (1:50K)
Regional	100k – 500k	ELF Regionalthemes
Global	> 500k	ELF Global themes

However , defining of common and well defined selection and resolution criteria for a specific LoD at European level is a real challenge. Up to now, there are no standard criteria for defining a specific scale (or LoD), each data producer having its own criteria.

Regarding Master Level 0 and 1, the ELF approach has been be to prioritize the existing most detailed LoDs of NMCAs without trying any harmonization between national criteria.

At Master Level 2, the ELF approach has been to set up common resolution and selection criteria based on what are the most commonly applied criteria among NMCAs. However the approach is not so obvious when looking to the discrepancies of the national thresholds declared for the same LoD.In a way, applying “generic rules” is not enough to achieve the level of harmonisation that is suitable for cross-border application.

At Regional and Global level, the ELF approach has been to keep the rules applied for EuroGeographics products[2]: EuroBoundaryMap (1:100k), EuroRegionalMap (1:250k) and EuroGlobalMap (1:1000k) and EuroDEM, whose resolution and selection criteria have already been well defined. The approach has been to meet European criteria by referring to already existing pan-European spatial dataset (i.ethe ECRINS dataset on watercourses and drainage basins produced by EEA). The approach has not been to purely apply the selection and resolution criteria of such dataset but to insure the data consistency.

The cross-border level implies that the data are consistent across-borders and properly edge matched on a boundary alignment like the international boundary.ELF defines three sub-levels for achieving the cross-border interoperability: the edge-matching process, the application of a European wide classification and the pan-European features.

The edge-matching concept consists of matching the features in their geometry and semantically between neighbouring countries and using as delineationagreed international boundaries. The concept is to set up connecting features (CFs) located on the international boundariesthat will serve as anchor points for the matching processing. The advantages of the connecting features are:

- Previously fixed edge-matching cases can be retained,
- It’s not necessary to have related data from neighbouring countries to process them to edge-matching,
- It prevents any discrepancies in an update cycle and availability of neighbouring datasets

The ELF project provide technical guidance on how to proceed and maintain edge-matching through data updates and tools to edge-match the spatial datasets at international

[2]EuroGeographics is the membership association of the European National Mapping, Cadastre and Land Registry Authorities (NMCAs), currently bringing together 60 organizations from 46 countries. http://www.eurogeographics.org

boundaries.The agreed international boundaries are collected into a European reference dataset (ELF IB) and made available on the ELF platform at all LoDs.

Providing geospatial data based on national classifications might not be sufficient for the work at the European level.Moreover, even if a European standard existed, another issue might come from the different way to interpret the classification scheme.

EuroGeographics has experienced such interoperability issue on the EuroRegionalMap(ERM) product for the road classification originated from national criteria. ERM has evaluated the application (mapping of national classification into ERM classification) for each country. Finally, recommendations for the mapping have been provided for several countries. The road classification has been improvedin ERM in order to establish a more harmonised and comparable European road network that fits for Basemap application and generalisation purpose.

Pan-European feature segments are identified as feature segments used as international boundaries or following them and shared betweenseveral nations like some watercourses. In the pan-European framework, those feature segments should get mutual consent for a single representation at all LoDs, which requires:

- Deciding of an identical geometry,
- Combining national attribute values whenever different,
- Populating a unique INSPIRE-ID. The option taken by the ELF project is to create a new one, instead of combining the national ones. Guidelines to populate INSPIRE-ID at European level have been defined.

This is the upmost level insuring full data interoperability and harmonisation in a pan-European framework.

Impact of the users

Setting up European data specifications is usually taken according two approaches: the top-down approach driven by the users and the bottom-up approach driven by the data producers.

Whatever the taken approach and the driven party, producers and users should decide of mutual consent for setting up European data specifications , which should be based on already existing data and which should both, meet the user needs and be acceptable for producers in time and cost for data collection and for reengineering national datasets according to the European specifications.

In that perspective, producers and users must cooperate. Providing reliable and authoritative national datasets that fit to European needs and setting up European data specifications accordingly should be a matter of a negotiated process between producers and users in a continuous dialogue.

The approachused for the EuroRegionalMap(ERM) product is a step by step process. Starting with what the data producers could produce in common, the dataset has been improved graduallyin level of interoperability based on users' recommendations. The producers and users agreed a programme addressing the priorities for improving the harmonisation of the data in an acceptable time and cost for reengineering the datasets. The datasets were regularly updated and improved.

The lessons learnt are that:

- The data producers are not able to improve the quality and fit for use of their data without a regular feedback of the users. A user feedback at each update or at least every year is required.
- The process may be time-consuming (8 years for ERM) but ensures a sustainable maintenance of the pan-European dataset originated from national authoritative data sources and allows to progressively adapt the national criteria to the European ones at reasonable costs for reengineering.

Semantic Interoperability for the ELF datasets

It is a general objective that the ELF datasets go beyond INSPIRE compliance in the matter of data harmonisation. It is important to estimate which degree of interoperability will be (or must be) achieved on the ELF data in the short, medium and long term.

The short term fits with the end of the ELF project.

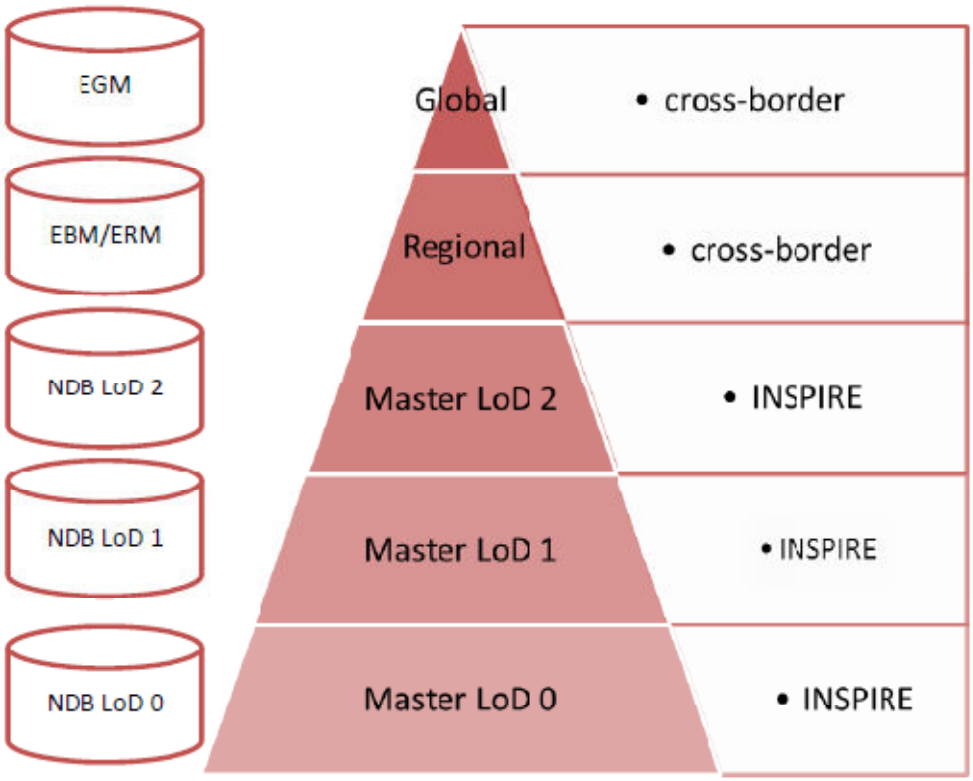

Figure 2: Semantic interoperability at different LoDs, in short term.

In the first release , the data producers will provide their currently available datasets,at Master LoDaccording to INSPIRE. At Regional and Global LoD, the data will be the current EuroGeographics' pan-European products EuroBoundaryMap(EBM),

EuroRegionalMap(ERM) and EuroGlobalMap(EGM) transformed according to the ELF data Specifications[3].

The medium term can be seen as two years after completion of the project when the ELF data specifications, and the geo-processing tools used for data quality, edge-matching, generalization, transformation, and change detectionall the outcomes of the ELF project will become operable and applied.

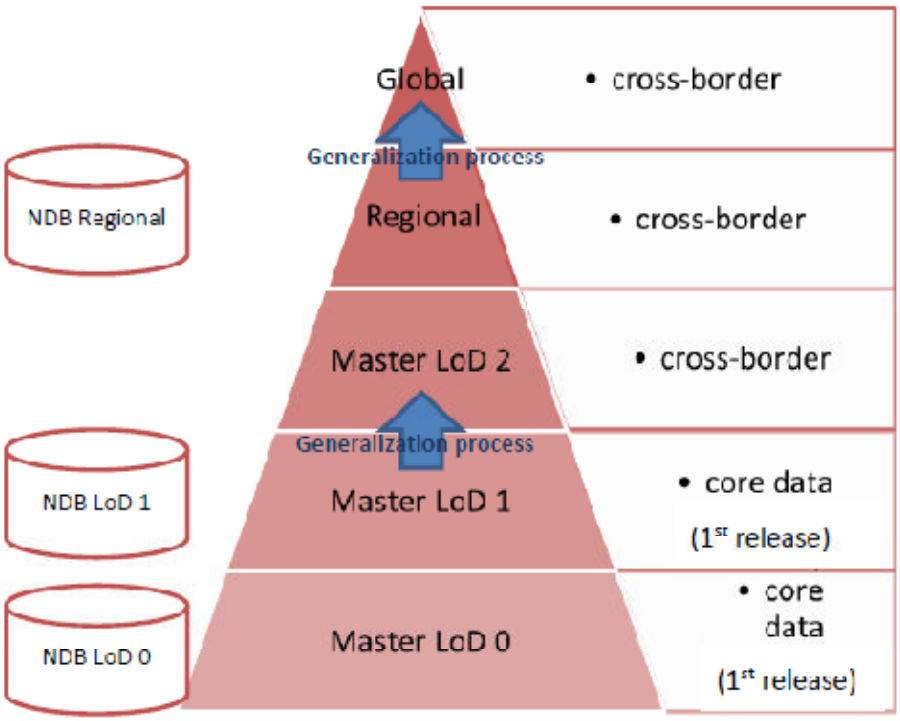

Figure 3: Semantic interoperability at different LoDs, in medium term.

At Master LoD 0 and 1, a core data content will be defined by the feedback of users' requirement and shall fit with what is needed for the generalization at Master level 2.

Master LoD 2 will reach a "well defined LoD", and tools for generalization between Master LoD 1 and LoD 2 will be operable.

The spatial datasets at Global LoD are semi-automatically derived from the Regional LoD.

Acquiring and collecting core data content at Master LoD 0 and 1 will take a "number" of years. This is the prerequisite to obtain a cross-border dataset at Master LoD 2, which will serve as reference data in future for deriving the Regional and Global LoD.

[3]ELF specifications : INSPIRE extended data specifications

The long term is based upon evolving user requirements, where the ELF data will need to meet the user's applications.

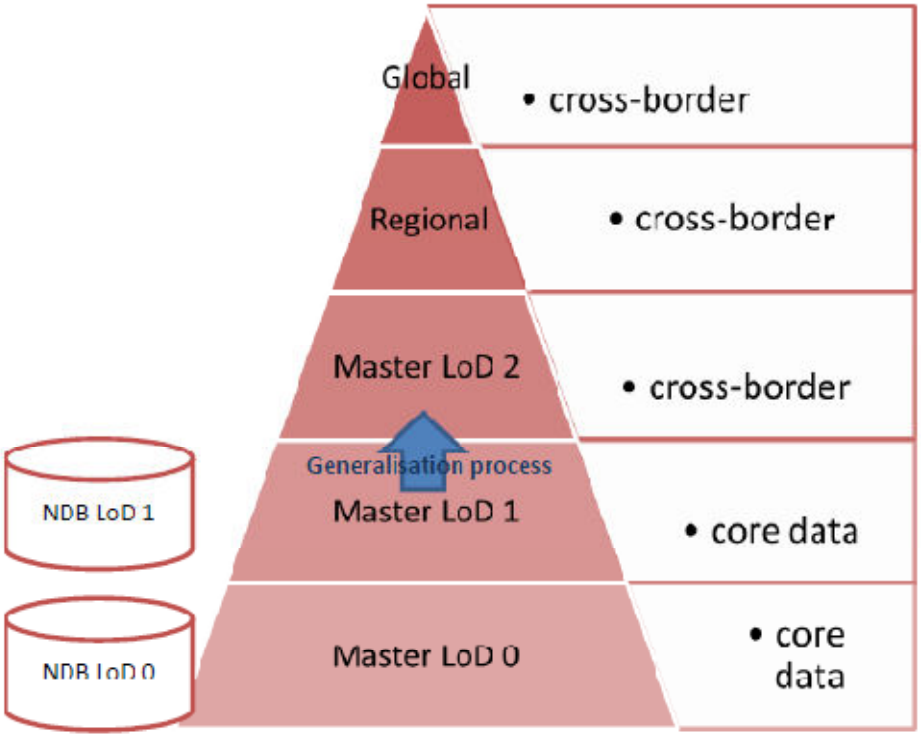

Figure4: Semantic interoperability at different LoDs, in longer term.

All suitable information should be collected and implemented at the Master LoD 0 and 1 , which provide the core content,required by most applications and available on a full European coverage.Based on the trusted core data content available at Master LoD 0 and 1, it will be possible to derive a generalized dataset at Master LoD 2, cross-border, and fitting the requirements for most of the applications requiring a European coverage.The Regional and Global LoDs will be derived from Master LoD 2 by generalization process.

However, reaching cross-border interoperability at Master LoD 0 and 1 for a European coverage may prove to be unrealistic, even in the longer term, owing to strong and irreconcilable national constraints in data collection and update and excessive costs for reengineering the data according to European mutual consent.

Technical interoperability

The paper will be restricted to the technical interoperability regarding ELF data processing and tools and not regarding the technical architecture and services.

The technical interoperability shall ensure proper ELF data workflow starting from national data sources to their publication in Web services and including the proper use of the geo-processing tools in quality, edge-matching, change detection and generalization. This process is referred to as the data supply chain.

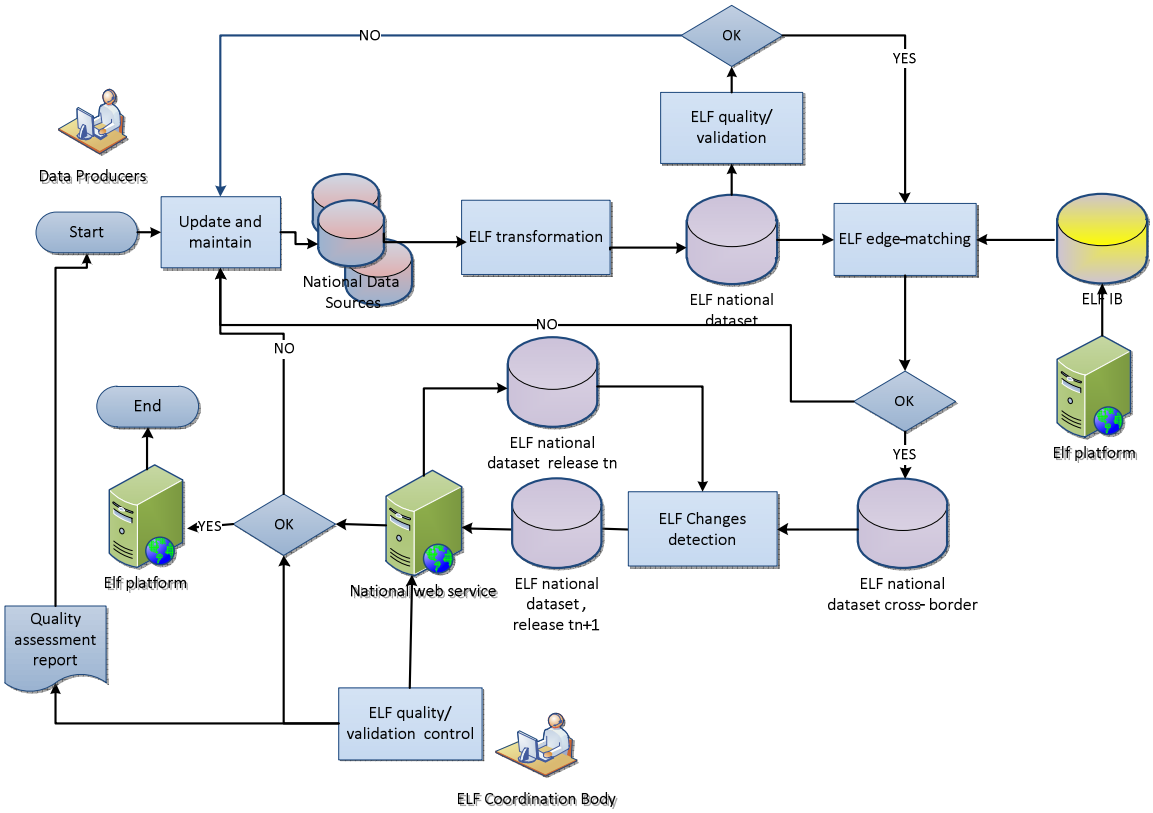

Figure 5: Data supply chain, cross-border

ELF transformation process: In principle, the geo-processing tools are implemented for running on ELF compliant data.

ELF Quality/Validation process: The process validates the compliance of the ELF data with the ELF data specifications and schema. This process can run at any time in case of update or correction on the ELF data. The process, at the end, provides metadata on the data quality like a reporting on ELF compliance and statistics on data content.

ELF edge-matching process: To achieve the cross-border level of interoperability, the ELF national component must be edge-matched properly on the ELF International Boundaries (ELF IB), by following the guidelines on the edge-matching procedure and using the geo-tool. The ELF IB database shall be downloadable from the ELF platform.

ELF change detection process: This process regulatesthe sustainable maintenance of the INSPIRE-Id and lifecycle information by comparison between two releases (Tn, Tn+1) for the same datasetand detects the changes in data updates which have occurred betweentwo releases. In tracking updates, the change detection process should be run as the last step, when no more correction or modification is brought on the ELF data.

To ensure a good quality maintenance of the data supply, a coordination body will proceed to a final quality/validation of the ELF data.

Defining a standard data format for exchanging data is a key condition for ensuring the technical interoperability in matter of geo processing and publication at European level.The technical interoperability principles adopted by ELF are the following:

- The geo-processing tools shall be able to run on ELF/INSPIRE compliant data sets.
- ELF/INSPIRE GML will be used as the standard for data exchange on the ELF platform.

Several tool developers are involved in the implementation of the geo-processing tools:

Tools	Tooldevelopers	Used Software
Data quality Validation	ESRI	ArcGIS
	1Spatial	1Spatial Cloud
	Delft University	prepair and pprepair
Change Detection	IGNF	C++ libraries
Edge-Matching	ESRI	ArcGIS
	1Spatial	Local installation of 1integrate with ELF Edge Matching Rules
	Delft University	prepair and pprepair
Generalization (Regional-Global)	IGNF	C++ programming based on IGN-F internal libraries
Generalization (master LoD1-master LoD2)	1Spatial	Local installation of 1Generalise with specific Flowline
	Delft University	tGAP builder (prototype implemented in Python)
	KadasterNL	ESRI ArcGIS
Transformation	Snowflake	GO Loader and GO Publisher

A survey inquiring on tools implementation and possible constraints for applying the ELF technical interoperability principles has been set up with the following questions:

- exchange data formats currently used for input and output
- the internal data format used for geo-processing
- are tools compliant to ELF specifications, and/or other data schema and specifications?

In summary, the internal data format is customized to each tools developer, the most commonly used exchange data formats are Shapefiles and SQL/PostgreSQL, Oracle.

All tools will be able to run on ELF compliant data sets (but are not restricted to). But currently no tools are able to work directly on GML data format. ELF/INSPIRE GML.Data have to be imported to workable data format used for the geo-processing.

After implementation and testing, some findings are:

- The ELF/INSPIRE GML data modelling is a little bit too advanced to be easily integrated into most of the GIS software and tools. Nowadays, the GIS world effectively uses flattened data structure, style relational data bases and import INSPIRE GML data into their workable data physical format. The issue is not so related to the GML format as such but well on the high-level elaboration of the INSPIRE data modelling implying additional implementation aspects that are not so easy to reflect from a physical data structure in a “relational” database.
- The geo-processing tools implicating modifications of the data like generalization, changes-detection and edge-matching will be restricted to simple features.
- The resulting file size is not so a big problem, but slows down the geo-processing when working on pan-European coverage or on a national coverage and loses efficiency on looping process like the validation/quality process.
- The INSPIRE UML data schema and modelling are great to represent the information in an elegant way, but are sometime disconnected from the intended use of the data like providing elaborated web mapping application.

Main Findings on achieving data interoperability

In the light of the best practices and experiences in the matter of pan-European production workflow, ELF has set up a stepwise approach for achieving semantic data interoperability with the aim to reach ELF cross-border datasets able to meet users’ needs and applications.

However, the cross-border level of interoperability providing pan-European coverage will not be achievable even in longer term at Master LoD 0 and 1, mainly owing to data availability and the national constraints for collecting and updating the data, but well for derived datasets at a lower level of details.

The ELF principle of data interoperability fully support the INSPIRE principles, in the idea that the INSPIRE/GML will be used as a standard for data exchange and publication. However, their application in the matter of geo-processinghas revealed some difficulties, which could lead to the risk of being rejected by the data producers.

- Paper Reference No.: **: PN-317**
- Title of the paper : **The European Location Framework (ELF), interoperability solutions for European public administrations and the INSPIRE Directive. Major challenges and lessons learnt for delivering authoritative, interoperable, cross-border geospatial reference data for a European coverage"**
- Name of the Presenter : Nathalie Delattre
- Author (s) Affiliation : National Geographic Institute of Belgium
- Mailing Address : 13, abbaye de la Cambre, Brussels, Belgium
- Email Address : Nathalie.delattre@ngi.be
- Telephone number (s) : +32 2 6298417
- Fax number (s) :
- Author(s) Photograph:

Brief Biography (100 words): *Nathalie Delattre works at the Mapping Agency of Belgium (NGIB) and has been involved in GIS and Operational projects management for 27 years. Since 2001, she has been leading the European production of "EuroRegionalMap" a pan-European geographical dataset distributed by the EuroGeographics association. She has participated in the European project "European spatial Data Infrastructure Network"(ESDIN) (2009-2012) and is contributing to the "European Location Framework" project (ELF) . Her main expertise is in geographical data production and maintenance within pan-European and INSPIRE frameworks. She is now involved in the INSPIRE programme at national level.*