
GS Soil – A user-centric approach for developing INSPIRE compliant geodata services for European soil data

Hermann KLUG, Bernhard BRETZ, Katharina FEIDEN, Carlos FIGUEIREDO, Sigbert HUBER, Rainer BARITZ, Endre DOBOS

1 Introduction

The European Directive 2007/2/EC for establishing an Infrastructure for Spatial Information in the European Community (INSPIRE) asks all EU member states to make interoperable data about the environment publicly available. The INSPIRE Annex III – Theme "Soil" is found to be one of the most heterogeneous and ambitious topics because of the diversity of products and terminologies used in this domain.

The GS Soil project "Assessment and strategic development of INSPIRE compliant Geodata Services" is aiming at establishing a European Soil Data Infrastructure. In order to tailor the results to the needs of users, a focus was laid on a high degree of user interaction and feedback.

This paper summarizes the most important outcomes of the project, which has led to many innovations in a broad context of geospatial information integration and dissemination.

2 Challenges and achievements

Every country or even every single data providing institution has its own soil data schema, classification scheme, legend, and laboratory assessment technologies. In addition, user requirements to soil information are originating from a great variety of sources (policy, research, industry, land owners) which has caused the same variety of products (soil basic maps and thematic maps in different scales and resolutions, local observations, monitoring). Therefore, there is a limited degree of match and harmonization along administrative boundaries. The diversity of content definitions and border effects is thus enormous.

The INSPIRE Directive defines the workflow to create interoperable data sets, which can be easily accessed based on well-defined metadata. Only very little Europe-wide environmental data sets are member-state-independent. Usually, member states contribute to the EU coverage with their own national databases. Due to the great diversity of soil data, it is often difficult to mosaic the national datasets and create an EU-wide coverage due to the lack of information on the data. This is the major driving force of the INSPIRE directive to create a framework, where all these national data sources could be used together as a joint dataset.

The main objective of the GS Soil project is to share datasets and best practices, improve and stimulate exploitation and the re-use of information on soils. The key successor for this

is the involvement of stakeholders, target users and soil data providers. The GS Soil group of soil and IT specialists together defined a framework of five major working components:

- Metadata to capture information about data,
- Interoperability of spatial data sets, attributes, legends, multilingualism and services,
- Network services (discovery, view, download, transform, invoke),
- Data and service sharing (policy),
- Coordination and measures for monitoring and reporting.

2.1 Developing a comprehensive metadata profile

Metadata is the information on the data itself. It describes how to interpret the data values, for example from which depth and date the data is coming from, who is the maintaining institution and how to get the data.

Soil data use requires much background information besides the values or classes that an object represents. Europe represents great variability of soil data with different approaches of field and lab work, mapping procedure and database development. Despite of the scientific and technical diversity, the soil datasets all over the world can be classified into typical classes and different categories within the classes. A simplified terminology and classification for data description have been developed for helping the users. Standardization of the data description - metadata development - is the only way towards interoperability. Since such metadata should be compatible and usable in a community and trans-boundary context, it is necessary to lay down rules concerning the metadata used to describe the spatial datasets and services. The major achievement of the GS Soil project is the development of a soil specific metadata profile for describing soil datasets as well as services and a Best Practice Guideline for helping the users to fill metadata of their contents according to the developed soil metadata profile.

2.2 Harmonisation and semantic interoperability

Investigation on harmonization and semantic interoperability of soil data requires core technical and content-related requirements so that different data layers and topics can be accessed in common portrayals and content services which can be understood and interpreted by users. With regard to INSPIRE, this does not necessarily mean full semantic harmonization, but requires the exchange of data in a commonly structured manner, a selection of core parameters, and the (optional) implementation of a content framework. Both aspects, the technical level of data exchange (GS Soil application schema) as well as best practice recommendations for harmonizing the content of soil data has been developed. Harmonization methods and the requirements to implement them are investigated in so-called test cases. The GS Soil test cases represent the whole spectrum of soil data addressed in the INSPIRE Definition of Annex Themes and Scope. A main focus is on the harmonization of soil maps in the scale of 1:250.000, because this resolution has been identified as a possible new high-resolution soil map of Europe. While the generic application schema for soil data can be seen as the backbone for soil data interoperability,

the test cases represent the project's core activity targeted to investigate and understand the requirements of soil data exchange under INSPIRE.

2.3 Technical interoperability

As there was the need for data providers to deal with interoperability of their technical systems, reference material has been developed which helps soil data providers to understand the purely IT-based requirements for interoperability. This is important to understand the need for harmonization, and to find a common language among data providers. While harmonized soil profile data were also provided when developing the European Soil Database 1:1 Mio, the documentation of harmonization methods applied to the respective national data bases is quite limited. Therefore, the source and quality of the data is not well-documented and understood. In addition, existing higher resolution soil maps are semantically heterogeneous despite a common mapping manual. The reason, the differences between existing products, and possible solutions with regard to harmonization are being developed as a new harmonization framework for Europe.

Major achievements within the technical framework that can be claimed:

- Creation of a working soil service network composed of a centralized portal and metadata catalogue and a connected and decentralised set of soil related catalogue and mapping services across Europe, with available and published soil related content.
- The collection and customization of a package of Open Software Tools directed to the data providers. These allow, in a common way, the publication of their soil metadata and datasets into the (GS)Soil network. The Open Software Tools package is composed of a Catalogue Server, a Mapping Server, a Transformation service and a Rights Management service.
- INSPIRE compliancy follows the latest documentation from the INSPIRE working groups, availability of soil data, plan of the project and the available GeoFOSS software packages. However, at this moment, not all available GeoFOSS software packages have implemented all the INSPIRE rules and/or have specific incompatibilities (e.g.: GeoNetwork schematron metadata validation).
- A comprehensive technical documentation, including the development of "cook books" (with logged partners' experiences), covering several aspects of publishing data providers soil data and metadata using the collection of Open Tools selected for the project (i.e. installation, configuration and even tips of adding data to the services). Similar auxiliary means is available for the centralised portal and metadata editor.

2.4 Usability and effectiveness tests

In summer 2011 a first testing phase regarding usability and effectiveness of the GS Soil Portal prototype was accomplished with more than 60 external users. Specific target groups were identified via questionnaires among project partners. Their results led to the definition of different user groups: "Commercial and Professional", "Utility and Public services", "Government and Administrators", "Research and Education", "Citizens & Public" and "NGO's and Non-profit organisations". Three testing scenarios were developed addressing the topics "search", "advanced search" and "maps" which consider also the usability of the

services of the GS Soil Portal. Testing participants have provided valuable feedback during the phase of testing. They indicated that the main goals of the portal, searching for specific data in a complex framework have been achieved already, also recommending some improvements towards better usability of the portal which will be considered in the final version of the portal. Such user feedbacks are of great importance in order to meet the expectations of the target users and serve their needs.

3 Summary & Outlook

The GS Soil Project has achieved a remarkable number of practical results towards a European Data Infrastructure for soil data. These achievements are even more valuable as they are optimized towards the specific needs of rather heterogeneous user communities. This also takes the aspect of usability very much into account, which is frequently neglected within the scientific community.

At the time of publication the project will be finished already (June 2012). At the moment following outcomes are seen as being highly innovative even beyond the domain of soil data infrastructures:

- **Metadata:** Profiles have been developed for soil datasets as well as data services. The decision on the suitability was based on the discussion and the agreement of a considerable number of project partners; amongst them also data providers. Best practice guidelines are a valuable mean to thoroughly support data providers in making their data sets INSPIRE compatible.
- **Semantic Harmonization:** Unique efforts in terms of scope have been made in order to address the huge harmonization needs. Several cross-border test cases have been intensively looked at and recommendations for data harmonization are provided.
- **Portal:** The work on technical harmonization finally lead to a one-stop portal (<http://gssoil-portal.eu/>), delivering data sets, services and further soil related information material. Most of the tools used for the Portal are Open Source, allowing individual adaptations where needed. Comprehensive documentation and cook books available under www.gssoil.eu, including user feedbacks have been developed to support people to be able to work with the portal in an easy way.

The GS Soil Portal is going to remain operative at least until 2014. A decision on either a prolongation or integration into an existing European Soil Data infrastructure is due then.

Acknowledgements

This project is co-funded by the community programme eContentplus under the contract number ECP 318004.