

The HABITATS Approach to Build the INSPIRE Infrastructure

Karel Charvát, Otakar Čerba, Štěpán Kafka, Tomáš Mildorf, and Přemysl Vohnout

Help Service - Remote Sensing, Benešov, Czech Republic
kch@hsrs.cz, ota.cerba@gmail.com, kafka@email.cz,
{mildorf, vohnout}@ccss.cz

Abstract. Different initiatives focused on spatial data in Europe should not be isolated but closely connected. The portfolio of such initiatives is very extensive. On the one hand, it covers all European activities such as INSPIRE or GMES, on the other hand, there are products of modern approaches based on neogeography and Volunteered Geographic Information (e.g. OpenStreetMap). Data are published by various regional or local authorities, non-governmental organisations, public bodies, research projects as well as by different commercial subjects. The focus of the HABITATS project was to build an environment that enables to share and combine data in order to reach new data, information and knowledge. On the basis of different pilots, HABITATS defined and tested harmonisation rules for spatial environmental data and designed the concept of Reference Laboratory as a tool for testing the interoperability and supporting unification of outputs cross different pilots.

Keywords: HABITATS, reference laboratory, INSPIRE, SDI, environment, social validation, volunteered geographic information.

1 Introduction

Advances in information and communication technologies contributed to increase the awareness about environmental protection all over the world over the last decades. More attention and is paid to the environment and changes towards the environment protection are made in policies, organisational structures, data collection and data management.

The Europe's nature conservation policy including the Birds and Habitats Directives was complemented by the Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE). INSPIRE supports environmental policies and policies or activities which may have impact on the environment. INSPIRE, which is based on spatial data collected by Member States (national spatial data infrastructures, SDI), paves the way for seamless data sharing and data reuse in Europe. Next to INSPIRE, there are other initiatives that support data sharing in different areas, such as Copernicus (previously known as Global Monitoring for Environment and Security) or the Shared Environmental Information System (SEIS).

This paper tackles the issues of spatial data management and SDI building based on the developments and results of the HABITATS project. HABITATS was an EU funded project focusing on adoption of the INSPIRE standards through a participatory process to design and validate environmental geospatial data, metadata and service specifications. HABITATS built an environment that enables combining data from various sources to gain new data, information and knowledge.

Spatial data in the context of the above mentioned initiatives help us to protect our lands and properties and the environment we live in and to shape our future environment. The attention is given to participation of citizens in decision making processes so they can influence where a factory or a new road will be built. Popular is collection of data using smart phones and other devices by volunteers. Citizens can for example report potholes in their street or fly tipping (dumping waste illegally) in their neighbourhood or map real world objects and phenomena.

The data sources for HABITATS were supported by the above mentioned initiatives as well as voluntary based projects such as OpenStreetMap. On the basis of different pilots, HABITATS defined and tested harmonisation rules for spatial environmental data and designed the concept of Reference Laboratory as a tool for testing the interoperability and supporting unification of outputs cross different pilots. The pilot applications tested ontology concepts that are described later in this paper.

2 HABITATS Networking Services and Architecture

The intention of the HABITATS Networking Services is to provide shift from the classic INSPIRE, GEOSS and GMESS architectures towards a solution which will support local and regional SDI building and their interaction with INSPIRE and which will move standard SDI model towards neogeography and the spider web paradigm of distributed systems. The way to test and provide it is Reference Laboratory as key tool of the HABITATS Networking Architecture.

The HABITATS Networking Architecture defines a platform-neutral SDI with a basic set of networking services in compliance with the INSPIRE Directive for sharing environmental data, especially that related to the 4 INSPIRE spatial data themes

- Sea-Regions (SR);
- Bio-geographical Regions (BR);
- HABITATS and Biotopes (HB);
- Species Distribution (SD).

This will result in a European metadata profile for these four data themes, which are an extension of the INSPIRE metadata profile. Our intention was not only to follow the INSPIRE profile for discovery services, but also to reflect the extension of the profiles for using data; a link to data modelling activities is therefore necessary. This profile is further open to extension by single countries or user groups, but the aim is that it is respected as a minimum set.

The set of HABITATS Networking Services was implemented on the HABITATS Reference Laboratory (RL) geoportal platform¹ introduced in the next section. This acts as a client of the 7 HABITATS pilots that provide a rich set of cross-pilots and inter-regional services.

In order to validate the HABITATS networking services architecture, a series of specific service applets were deployed and tested for data sharing using the HABITATS RL geoportal platform.

Comparing network services and spatial data services, the INSPIRE Forum² definitions are:

- Spatial data services are all services (Discovery, View, Download, Transformation, Invoke, Other) regarding spatial data;
- Network services (Discovery, View, Download, Transformation, Invoke):
 - non-compliant network services are INSPIRE compliant services with respect to functionality;
 - compliant network services are compliant services with respect to functionality and quality of service.

The HABITATS networking architecture supports INSPIRE Network services, but needs to go behind this concept. INSPIRE networking services are in principle limited only to the management of existing data and metadata. The HABITATS Networking Services also support such functionality with data and metadata management, data and metadata collection, working with non-spatial data, etc.

The HABITATS service applets re-use existing applications where possible and are themselves designed for re-use. The selection of the specific services to deploy is primarily a user-driven process, as defined in the user scenarios and requirements. The prototype set of Network Service Applets to be installed in validation pilot platforms were drafted as:

- A series of specific networking service applets deployed and tested for data sharing within the project using the Network Service Architecture:
 - Interoperability Services;
 - Enabling Services:
 - Visualisation of information layers;
 - Overlay of information from different sources;
 - Spatial and Temporal Analysis;
 - “quick” and “light” on-demand applets to meet validation pilot expectations and user needs;
 - Usability, simplicity and openness to rapid prototyping mash-ups;
- A set of specific service applets that allow users to identify, access, use and reuse habitats-related data, designed and deployed on-demand to meet user needs;

¹ At www.habitats.cz

² http://inspire-forum.jrc.ec.europa.eu/mod/groups/topicposts.php?topic=11135&group_guid=8651

- Mobile Apps allowing to use advantage of the HABITATS RL;
- “Quick Prototyping” service applets respecting the HABITATS service architecture and developed on-demand.

These are based on the outcomes from the earlier tasks and work with the HABITATS RL, and will now lead into the interface tools and toolkit.

The HABITATS Networking Architecture aims to extend the principles of the INSPIRE networking architecture, because INSPIRE doesn’t cover important aspects such as data management and data collection. So all components of the INSPIRE networking architecture will be included in the HABITATS architecture, but this concept will be extended by other functionalities. From this point of view principles of GEOSS and GMES and also principles of Shared Environmental Information System (SEIS) and Single Information Space in Europe for the Environment (SISE) have influenced the HABITATS architecture and its networked services.

The development of the network service architecture process was initiated through a state of the art analysis of existing SDI, to find out more about already existing infrastructures and to examine how data should be shared and what services are required to enable sharing. When designing the networking architecture, a set of specific networking service applets was deployed and tested for data sharing within the project. Also the potential for re-use of existing application was taken into account.

3 HABITATS Reference Laboratory

The HABITATS Reference Laboratory (RL) is the central hub of the HABITATS Networking Architecture. It consists of several layers, which include [2]:

- Data layers – management data and files on storage, eventually guarantee access to external sensors;
- Server (engine) layer – defines tools, which guarantee basic services on the server side – supplying service;
- Client layer – is client side of Web services, which guarantee access of users to services;
- Application layer - is some form of wrapping elementary client services into application or into such form, which could be used by other web tools;
- Presentation layer - contain such web tools, which allows to combine and to publish single objects from the application level as part of Web presentation.

Figure 1 shows the different layers of the HABITATS Networking Architecture.

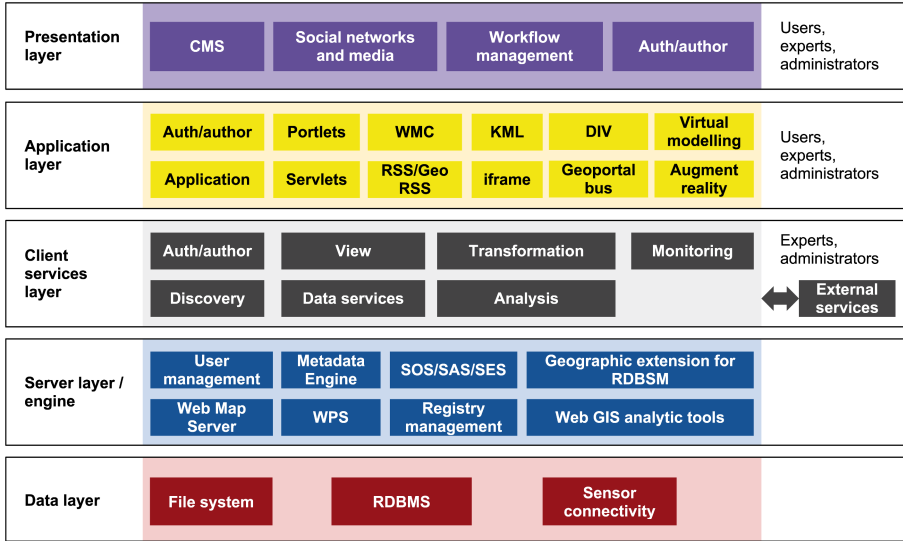


Fig. 1. HABITATS Networking Architecture [3]

The final implementation of the HABITATS services anticipates that selected concrete services will be deployed for every pilot, and that there will be one central platform (i.e. the Reference Laboratory).

The HABITATS RL provides the Networking Architecture supporting both Networking Services and Spatial Data services, that support the SDI network services to enable trans-European sharing of habitats-related spatial data between public authorities and other stakeholders in the Community, enabling the creation of value added services. The HABITATS RL is focused on implementation of:

- Cross pilot scenarios based on sharing of data among more pilots;
- Validation platform for testing of conformity of implemented pilot services;
- Services supporting global discovery, view and downloading;
- Repository for common metadata;
- Repository for pan-European datasets such as Natura 2000, Corine Land Cover, Urban Atlas and Open Street Maps;
- Interlink with social networks.

The HABITATS RL provides the Networking Architecture, that supports the SDI network services to enable trans-European sharing of habitats-related spatial data between public authorities and other stakeholders in the Community enabling the creation of value added services.

The HABITATS RL enables deployment of specific service applets, including interoperability and enabling services, on-demand from user communities and the pilots for initial implementation and validation. It is being developed further to include an invoking service toolkit integrating the service applets with the goal of facilitating the development of end-user services accessing habitats-related spatial data over time.

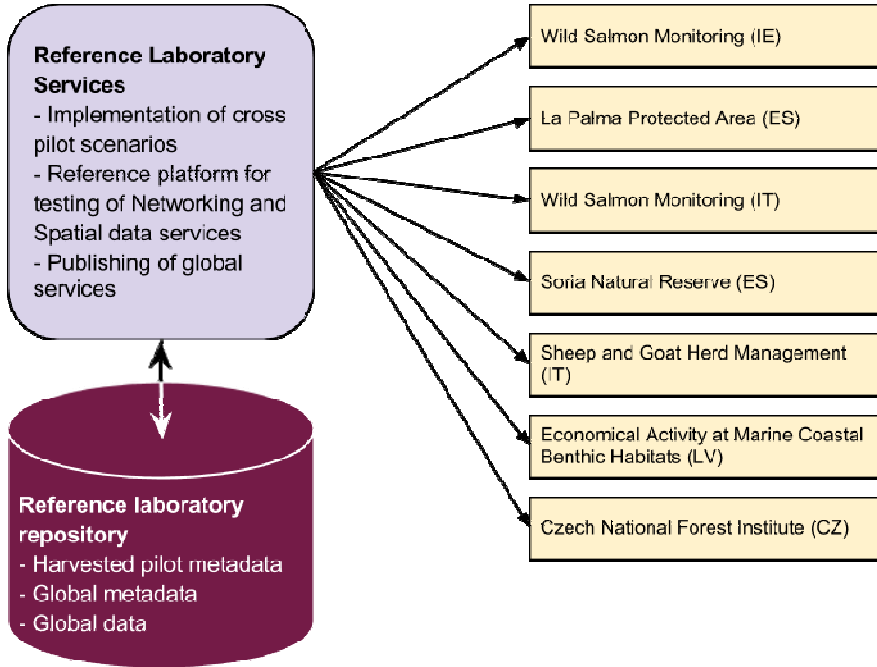


Fig. 2. Reference Laboratory and its pilot applications [3]

However, applications are the key objectives and final goal of using the HABITATS RL. As the HABITATS RL is just a geoportal tool to help to build applications that address the Pilots' typical use cases.

The HABITATS RL is:

- an interface that enables interactive search, portrayal, evaluation, sharing, analyse and reuse of spatial and non-spatial data;
- a solution based on interoperable standards (OGC, W3C, OASIS, ISO). It is interconnected to other resources through the Internet. It helps to create a distributed structure of information and knowledge with spatial position.

However, the RL is not a central data storage nor a closed web application with maps. It is a geoportal with:

- Independent components;
- Composition according to user requirements;
- Based on SOA;
- Possibility to integrate with other resources;
- Maximum openness;
- Open Source;
- Open Standards;

- Extension to non-GIS community;
- Open Search;
- Administration of other (non-spatial) data sources.

The HABITATS RL allows deployment of the current state of the art of technological solutions, which will be tested and adopted by the HABITATS partners and user stakeholders. It allows testing of current existing technology and generation of further research tasks driven by users. The HABITATS RL also collects information coming from other projects, which is an important input for the HABITATS analysis and public discussion. The methods of social assessment will be an increasingly important part of the HABITATS RL.

Thus the HABITATS RL's networking services aim to help HABITATS to extend user-centric, co-design approaches into the arena of standards design and adoption processes, considering standards initiatives such as INSPIRE, OGC, UNSDI to be significant social, economic and institutional innovations. The elements of the approach are maintained, applying the model at all levels from the global scale to the local and regional policies that frame many HABITATS validation pilots. Community building activities follow a Web 2.0 approach to capture the knowledge in active user communities with a strong interest in contributing to the standards development process. By inviting a broad multi-sectoral and inter-disciplinary range of concerned stakeholders to participate into the HABITATS network, a viral motivation spiral is set off. A peer-to-peer approach to opening up information sources and providing access to content ensures a rapid extension of the critical mass of environmental data established by project partners.

4 Ontology Concepts Applied by the HABITATS

4.1 Domain Ontology

Domain ontology describes the fundamental concepts and relationships which are important in a particular domain. It represents a domain vocabulary as well as a formal and formalized specification of the meaning of terms. According to the Protégé web site (<http://protege.stanford.edu/>) ontologies range “from taxonomies and classifications, database schemas, to fully axiomatised theories”. Ontologies could be an essential part of many applications of information technologies (including geoinformation technologies) such as portals, information management and integration systems or semantic web services.

The principles of ontologies could be summarized to the following three points:

1. Ontologies serve as a tools used for a description of real or abstract objects or processes;
2. Ontologies as a kind of data source have to use formal recommended to use and distinction between two words connected with ontologies – formal (application of formal terms) and formalised (formal way of coding) ways of data structures;

3. It is not possible to develop a universal ontology, since its properties depend on various factors such as way of use.

Common domain ontologies using the standard data format OWL (Web Ontology Language, see below) can contain five crucial elements or groups of elements [4]:

1. Classes represent and describe concepts (particular parts of domain). They are also able to interconnect these parts. They form together with individuals a hierarchy of ontology. The hierarchy is based on is-a (is a part) relation. There are two types of classes in ontologies – primitive classes (described by necessary conditions) and defined classes (described by necessary and sufficient conditions). To build ontological classes the formulation of descriptive logic (such as union, intersection or quantifiers) are used;
2. Individuals are very similar to classes. The indivisibility of individuals is the main difference between these two components. While class can contain other class or classes, the same construction cannot be implemented to individuals (individuals cannot be divided);
3. Properties or relations are designated for definition, interconnection, description and detail specification of classes, individuals and their connections. There are two basic types of relations in ontologies: DataType Properties (connecting classes or individuals with concrete data types) and Object Properties (describing relations between classes or individuals);
4. Axioms express explicitly rules managing relation between elements of ontology. They enable to describe rules such as equivalence, subsumption, disjunction etc. Ontologies use two specific axiomatic constructions – closure axiom (combination of existing and universal quantifier) and coverage axiom (definition of closed set of sub-classes);
5. Annotations represent an analogy to metadata records. They are based on Dublin Core standard (<http://dublincore.org/>) and enable not only to describe the whole ontology but also its elements by properties such as label or language.

The methodologies constitute very important but underestimated part of ontologies – designing of ontologies as well as development and managing phase. There are many various methods such as On-To-Knowledge, METHONTOLOGY (the most popular methodology; standard IEEE Std 1074-1995), Cyc (probably the oldest methodology), Gruninger and Fox, DILIGENT, KACTUS, SENSUS, Noy and Macguinness or Uschold & King. They are used to construct and maintain ontologies. It is necessary to mention that most of the methodologies are based on detail preparatory phase which is very often underestimated in practice. This phase can be composed of an identification of the ontology (title, purpose, scope, coding, granularity), planning of conceptualisation and an integration of existing ontologies and other sources.

4.2 General Example Using Ontology in a Pilot Application

Ontologies were used in selected pilot activities in the HABITATS project. They are not designed for spatial data harmonisation, but they should be a part of the harmoni-

sation process. They can serve as a description of various heterogeneities and mapping of various types of concepts. The real harmonisation is performed by harmonisation tool (GIS, ETL etc.) and using information and knowledge formalised and stored in ontology.

As an example, the ontological system transforming various types of taxonomies (classification systems) of land cover and land use data was used. Land cover and land use, widely used by the INSPIRE initiative, form an essential part of the HABITATS pilots (forestry, sea regions, tourism, etc.).

The ontology is composed of three main components:

1. File *Parameter.owl* that is imported into ontologies containing various classification systems, describes the properties of all classes of these nomenclatures. The file contains not only all the parameters (the current version consists of nine parameters), but also their values and relations (object properties) that assign values to the parameters of each class of classification systems. Current version covers nine properties with values arranged in multi-level hierarchy;
2. Files containing all elements of the particular classification systems (CORINE Land Cover, PELCOM, GlobCorine, LUCAS, UMD, NLCD etc.). All classes of ontologies are described through the properties using the parameters and their values;
3. Ontology is used to transform each other classification systems (file *Trasformation.owl*) – this ontology imports files describing particular classification systems. After their evaluation (reasoning) is obtained multi-hierarchical structure providing information about the interactions of described taxonomies.

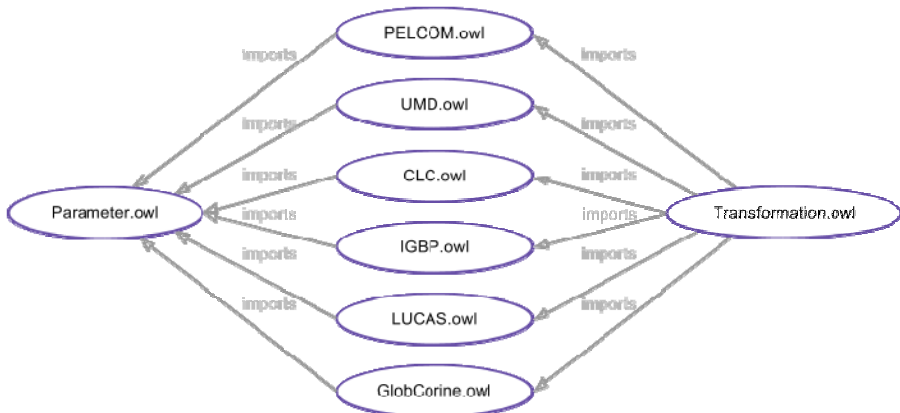


Fig. 3. Ontology relationships [1]

The explicit parts of ontology (Figure 3) are processed by reasoners. The inferred ontology contains relations between particular classes of taxonomies.

5 Conclusions

The HABITATS project built an environment that enables to share and combine data from various sources. The project results were validated through the HABITATS Reference Laboratory and pilot applications. On the basis of different pilots, HABITATS defined and tested harmonisation rules for spatial environmental data and designed the concept of Reference Laboratory as a tool for testing the interoperability and supporting unification of outputs cross different pilots.

References

1. Čepický, J., et al.: D3.3.2 Data models and transformation processes (2012)
2. Charvát, K., et al.: D4.2.2 HABITATS INSPIRE Networking Architecture (2011), <http://www.slideshare.net/kcharvat/habitats-d422-final>
3. Charvát, K., et al.: D4.3.2 HABITATS networking services and service toolkit (2013)
4. Čerba, O.: Ontology Of Cartographic Terminology. In: International Conference on Cartography and GIS (2012)