

HUMBOLDT Scenario: Protected Areas - Harmonisation Capabilities for a Future ESDI

Paolo VILLA^{a,1}, Roderic MOLINA^b, Mario A. GOMARASCA^a,
and Emanuele ROCCATAGLIATA^b

^a*CNR-IREA, Institute for Electromagnetic Sensing of the Environment,
Via Bassini 15, Milano, Italy (villa.p@irea.cnr.it)*

^b*GISIG, Geographical Information Systems International Group,
Via Piacenza 54, Genova, Italy*

Abstract. The HUMBOLDT project, started in October 2006 and supported by the European Community has the aim of implementing a Framework for harmonisation of data and services in geoinformation domain, under the INSPIRE Directive and in the context of GMES Initiative. The two-pronged approach of HUMBOLDT comprises a technical side of framework development and an application side of scenario testing and validation. As the outcomes of implemented harmonisation framework grows mature, the application side of the project grows more and more prominent. Among the HUMBOLDT application Scenarios designed to demonstrate the capabilities of the Framework there is the one covering Protected Areas themes and use cases. Protected Areas Scenario aims to transform geoinformation, managed by park authorities, into a seamless flow that combines multiple information sources from different governance levels (European, national, regional), and exploits this newly combined information for the purposes of planning, management and tourism promotion. The scenario defines a number of use cases to address and examine planning and management issues, as well as tourism added value. The scenario develops the harmonisation process via active engagement with various stakeholders at the national and trans-national levels including national authorities and European agencies. To test and create examples of the use of the HUMBOLDT tools a Desktop and Web GIS environment together with a server environment was set up and the resulting tests with two HUMBOLDT Web Processing Services have been documented.

Keywords. Data harmonisation, INSPIRE, ESDI, protected areas

Introduction

Harmonised geoinformation is a basic need for fulfilling the task of creating a Spatial data Infrastructure reliable and efficient, in which different data sources and different services for discovery, portrayal and retrieval of geodata is a crucial asset [1,2]

At the European level, the road to an ESDI follows the guidelines contained in the INSPIRE Directive of the EC; in this context, the structure of an ESDI shall be composed of a set of interoperable, interacting services, thus following the Service Oriented Architecture (SOA) paradigm. Such an architecture matches the distributed responsibilities regarding service provision and data management in the geoinformation sector well. For a SOA to work, an essential element is to select or build a group of interface standards that are mutually interoperable and complementary[3].

Within the scope of the HUMBOLDT project, after a first analysis of harmonization-related studies and efforts, the following working definition for data harmonization is taken into account:

Geo-data harmonization implies and means the possibility to combine data from heterogeneous sources into integrated, consistent and unambiguous information products, in a way that is of no

¹ Corresponding Author.

concern to the end-user. Geo-data harmonisation implies and means the possibility to combine data from heterogeneous sources into seamlessly integrated, consistent and unambiguous information products, in an easy and repeatable way, adapted to the end-user's requirements and context [4].

For better comprehension, it is useful to approach the different needs for data harmonization through a list of undistinguished causes of heterogeneity in spatial data[5]. Heterogeneity in the case of spatial (geographic, atmospheric, geological) data is for example caused by differences in:

- data format and data collection procedures
- spatial reference system
- data/conceptual model: structure and constraints
- metadata model
- nomenclature, classification, taxonomy, terminology, thesaurus, ontology
- scale, degree/amount of detail, extent (spatial, thematic, temporal)
- portrayal (legend/classification, style)
- processing functions: their parameters and formulas/algorithms.

1. Geodata harmonisation within the HUMBOLDT project

The four-year EU project HUMBOLDT contributes to the implementation of a European Spatial Data Infrastructure (ESDI) that integrates the diversity of spatial data available for a multitude of European organisations. The main goal of Humboldt is to enable organisations to document, publish and harmonise their spatial information. The software tools and processes created demonstrate the feasibility and advantages of the initiative INSPIRE (Infrastructure for Spatial Information in Europe) as planned, meeting the goals of Global Monitoring for Environment and Security (GMES).

Under the coordination of Fraunhofer Institute for Computer Graphics 28 project partners from 14 European Countries work together on 12 work packages. The HUMBOLDT project partners represent private companies, public authorities, universities and other research institutions, which prevents the project from taking any kind of narrow perception of spatial data and processes.

The work packages are organized along a two-pronged approach (see Figure 1) has been designed and followed. This approach focuses on integrating both concrete application requirements but also technical innovations, best practices and research results [6].

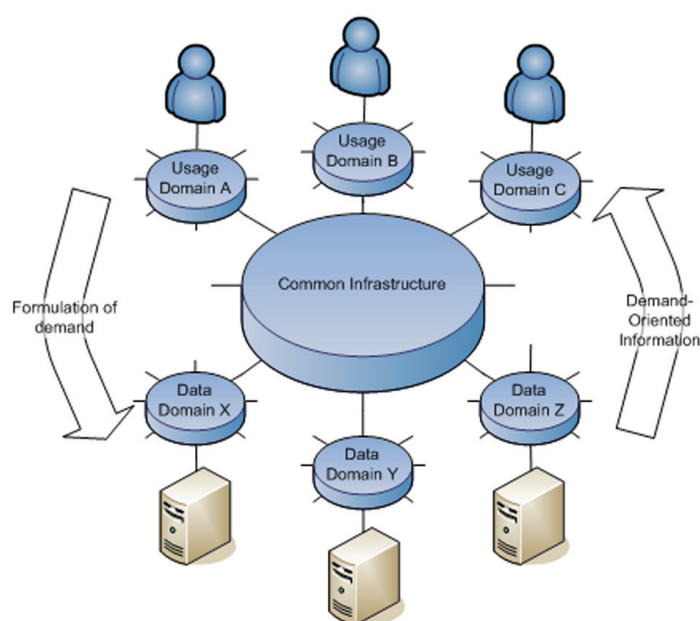


Figure 1. The user-driven approach of HUMBOLDT project to geodata harmonisation.

An essential element of the application driver side of the project is the development of so-called HUMBOLDT scenarios in which the different components are applied and tested under realistic conditions and which represent different GMES application fields, ranging from border security to urban planning, risk management and the protection of nature.

For these scenarios, a process analysis is done which shows the steps necessary to harmonise data and metadata. The different stakeholder groups have to act as drivers, to ensure that the technological development fulfils their requirements as good as possible.

Currently, the following application scenarios are being addressed by the Humboldt partnership. They are organized along the core themes defined for GMES:

- **Border Security:** Effective Border Control and Security in Rural Areas
- **Urban Planning:** European Urban Management Information Systems
- **Urban Atlas:** Enforcing GMES in Urban areas mapping core services
- **Forest:** Saxony & Czech Cross-Border Forest Scenario
- **Protected Areas:** Management of Protected Areas
- **ERiskA:** Environmental Risk
- **Transboundary Catchments:** Crossborder water basin Management
- **Ocean:** Oil/Contaminants spill crisis impact and management
- **Atmosphere:** Integration for Atmospheric Data Distribution

At the core of the development work done within the HUMBOLDT project stands the HUMBOLDT framework. This software framework is the hull for the various data harmonisation scenarios and provides the common functionality required by the scenarios. The framework itself as a project outcome mainly addresses the developer community and the application developers using it to create end-user applications and scenarios.

Currently, the second version of the HUMBOLDT Framework (see Figure 2) is available, together with a Grounding Catalogue and an initial set of data for the assessment of these prototype functionalities with special reference to application areas indicated by HUMBOLDT Scenarios.

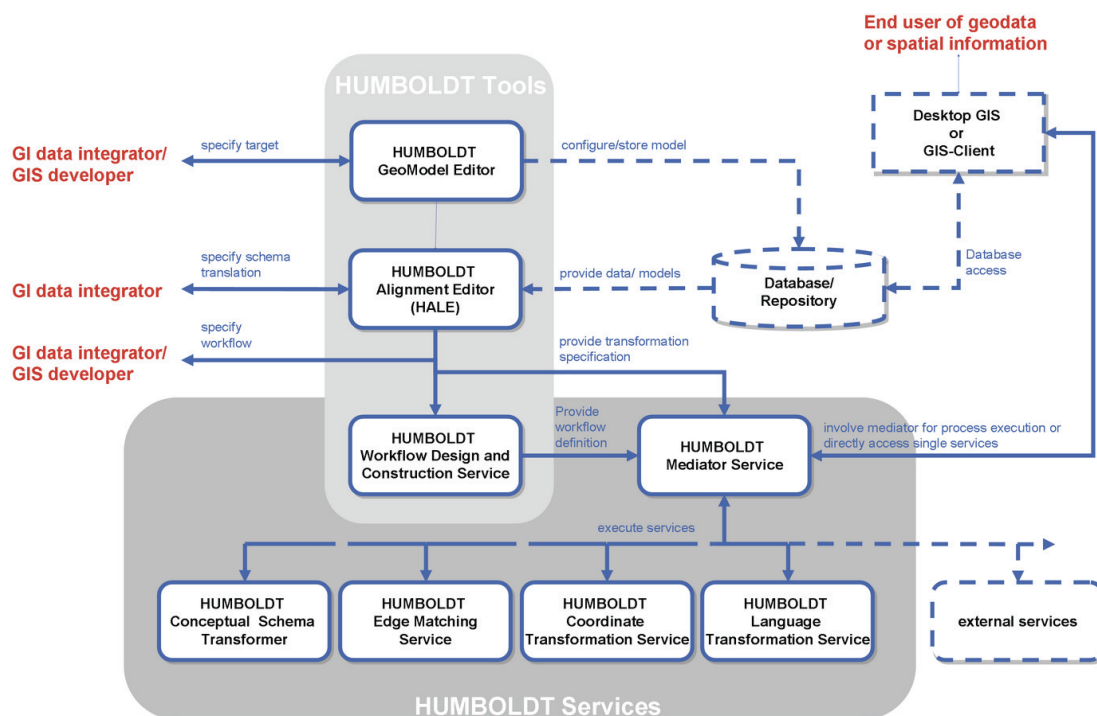


Figure 2. Overview of HUMBOLDT Framework Components.

The Architecture of the HUMBOLDT Framework [7] has been based on an approach which comprises as the fundamental part a Mediator Proxy able to support standard interfaces (Figure 2). More specialized interfaces are also integrated in the Framework Architecture, as:

- **Mediator Service Component:** This component orchestrates transformations of data sets and also provides the entry point to other services.
- **Model Repository Component:** This component can be used to store, manage and discover Application Schemas and Mappings between those.
- **Context Service Component:** The context Service allows users to define rules for the products they want to have at the end of a transformation process.
- **Information Grounding Service Component:** This component provides concrete services matching the user's requirements.
- **Workflow Service Component:** The workflow service automatically discovers heterogeneity issues and applies designs a workflow for the resolution of those, according to a user's request and context.

Alongside with other tools and services such as:

- A **Workflow Design and Construction Service** (WDCS);
- An **Alignment Editor** (HALE);
- A **Geomodel Editor**;
- A set of **Transformer Services** (edge matching service, language transformer service, coordinate transformation service, ...).

The transformation and harmonisation process and its feasibility and efficiency, strongly depends on the availability of the description of transformation rules from the viewpoint of the conceptual schema level.

All those aspects and issues are being tackled within the HUMBOLDT project environment, on the path leading to the actualization of INSPIRE efforts into an ESDI.

2. An application scenario: Protected areas

Among the HUMBOLDT application Scenarios designed to demonstrate the capabilities of the Framework there is the one covering Protected Areas themes and use cases. Coordinated by GISIG, the Protected Areas Scenario aims to transform geoinformation, managed by park authorities, into a seamless flow that combines multiple information sources from different governance levels (European, national, regional), and exploits this newly combined information for the purposes of planning, management and tourism promotion [8].

The Protected Areas Scenario Demonstrator Portal has been developed and the first version of one of the foreseen Application Cases is described in the work. To test and create examples of the use of the HUMBOLDT tools a Desktop and Web GIS environment together with a server environment was set up and the resulting tests with two HUMBOLDT Web Processing Services have been documented.

One of the main objectives of the HUMBOLDT project is to provide tools to map and transform complex database and application schemas. In this sense, the current work of the Protected Areas scenario is focused in harmonizing Protected Areas data from various countries using the HUMBOLDT Alignment Editor. The aim of this tests is the creation of step-by-step examples on how to use the HALE tool covering different aspects of schema mapping and different functions in terms attribute transformation, giving special relevance to testing and exploit the Annex I Protected Sites data theme for INSPIRE-compliant data provision.

All the operations have been done using open source tools, from the pre-processing of the data sets to the visualisation of final results.

The scenario defines a number of use cases for which detailed user stories have been developed to address and examine planning and management issues, as well as tourism added value. The scenario develops the harmonisation process via active engagement with various stakeholders at the national and trans-national levels including national authorities and European agencies. A schema of the HUMBOLDT components architecture which are deployed for Protected areas Scenario, so far, is given in Figure 3.

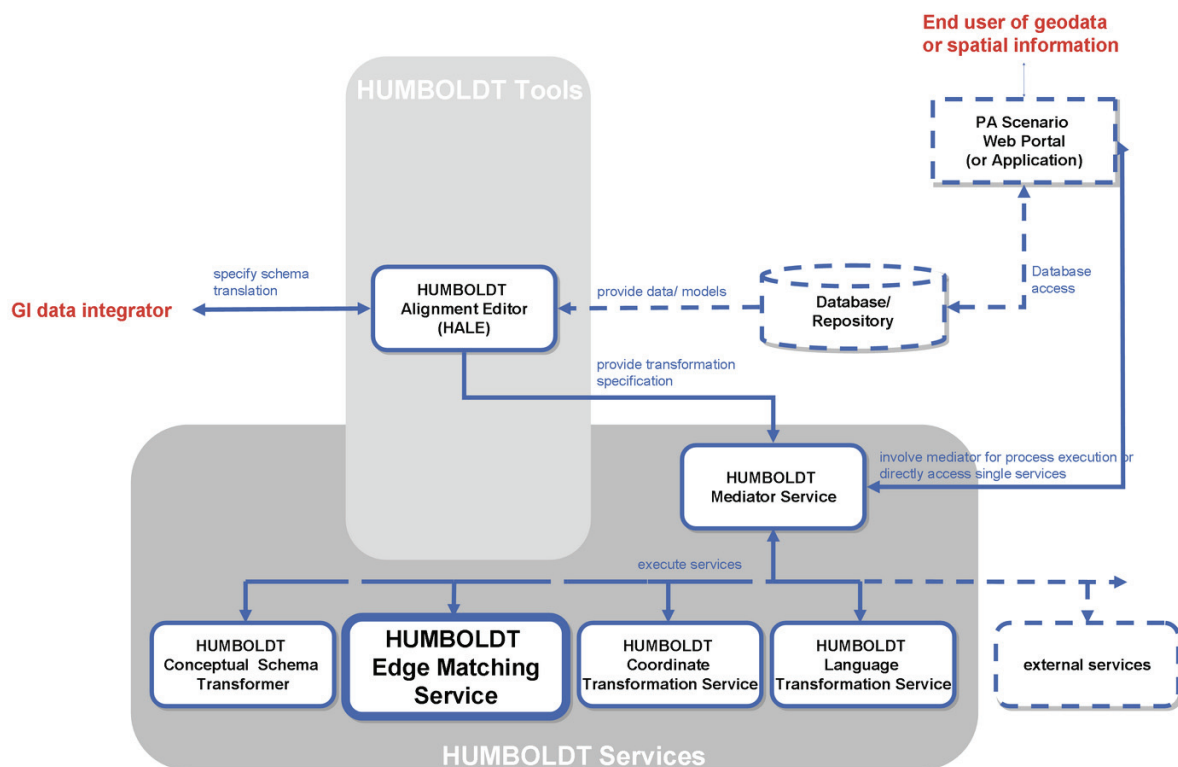


Figure 3. Overview of HUMBOLDT Framework Components architecture, as utilized in Protected Areas Scenario (highlighting the principal transformation service utilized: Edge Matching Service).

More specifically the use cases investigated by the scenario are related to a strategy of which the pursued aims are:

- Creation of a geo-spatial repository where stakeholders contribute and share available geoinformation from any source; final users can browse this information.
- Management (use of geo-information by planners and officers)
- Promotion of its assets for a sustainable use (access to geo-information by citizens in order to receive their useful feedback)

Namely, the HS Protected Areas is investigated in order to provide harmonisation support especially for the interaction between various levels of work and administration: management bodies, local stakeholders, national authorities, European agencies, cross-border administrative bodies.

The scenario is tested in two areas, one between Portugal and Spain, covering the Douro river natural Park in Portugal and the Arribes del Duero Park in Spain, and in Italy, covering the Beigua Regional Park. An example of harmonisation using HUMBOLDT Edge Matching Service in the context of Protected areas Scenario is given in Figure 4.

The below scheme of actors is drawn upon the following classification of purposes in using geodata and geo-information:

- Preservation
- Re-naturalisation (for example in changing situations such as landscape change and protected landscape)

- Exploitation (for sustainable development)
- Commercial use (e.g. for tourism)
- Science and education

Beside the scope of use, you can consider the kind of target users, an example of which are the end users, involved in browsing geo-information (aggregated information: the lowest level of access) or geo-data (information elements). They are decision makers, tourism operators or citizens (the last ones intended as persons). In Table 2 the classes of users are summarized.

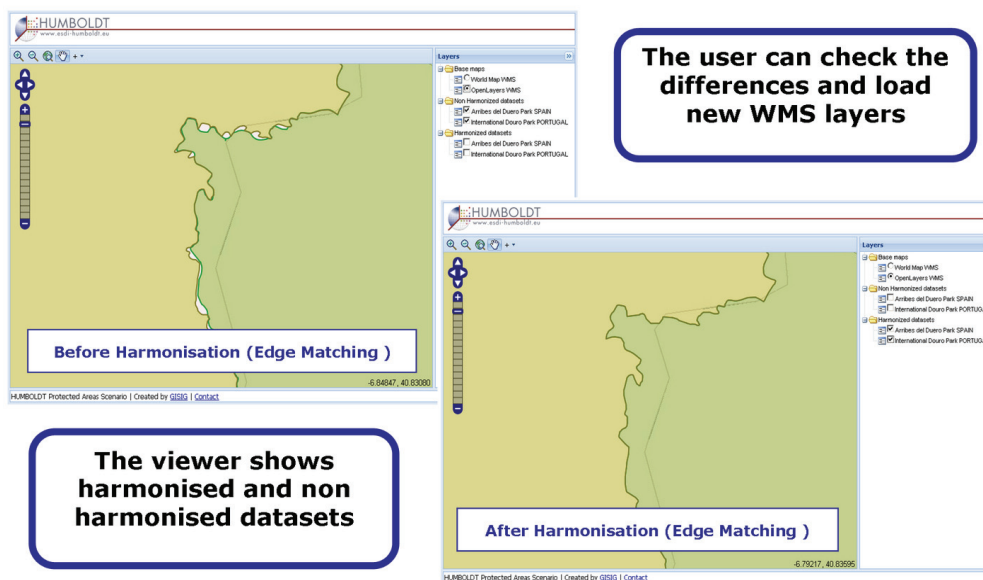


Figure 4. Example of Edge Matching Service WPS results over cross-border (Spain and Portugal) Protected Areas data layers.

Table 2. Classification of geo-information users and characteristics in relation to the Humboldt framework

	HUMBOLDT framework Users	Scenario only Users	Type of client	Users in the Protected Areas Scenario
End user (geodata)	No	Yes	Thin	Decision makers Leisure and tourism operators
End user (geo-information)	No	Yes	Thin	Citizens (persons)
Data integrator	Yes	No	Thick	Planners and officers Citizens (Associations and NGOs) Researchers
Data provider	Yes	Not relevant	Not relevant	Commercial operators: Geo-data providers Commercial operators: Geo-services providers
Data custodian	Yes	No	Not relevant	Mapping agencies Territorial governments in charge, e.g. regional

As already mentioned, Protected Area Scenario is structured into Use Cases that consist in applications of different harmonisational instances related to different requirements and users. A brief summary of the considered use cases is given here (of which the the first two will be developed in detail):

- **Management of a Protected Area** - The main task is to create plans and managing the protected area.

- **Tourism valorisation in a Protected Area** - The main task is to exploit at the best the area and enjoy its offer of nature
- **Impacts of infrastructure on a Protected Area** - The main task is to communicate the impact of infrastructure on protected areas.
- **Data processing in a Protected Area** - The main task is enabling local authorities to perform environmental analysis on protected areas geodata.

3. Summary and outlook

This paper has described the structure and approach of the HUMBOLDT project, giving a rationale for the HUMBOLDT framework capabilities and discussing the rationale and outcomes of HUMBOLDT Application Scenario Protected Areas.

The major aim of HUMBOLDT is the implementation of efficient, cost-effective, reliable, generic, interoperable and sustainable solutions for the issue of spatial data harmonisation and integration of geographic services in the framework of an ESDI. This objective is to be reached by putting INSPIRE principles into practice, applying international standards and using as core reference the users' requirements and needs, finally establishing a community of users and developers.

HUMBOLDT framework is an architecture of software components and services aimed at managing the harmonisation process of geoinformation within the European context. The methodology of the HUMBOLDT development is based on a dual approach, comprising both a technological and an application side, and on an iterative process of implementation, during which the solutions found are tested and validated with the cooperation of an application momentum, composed of scenarios which cover topics of utmost importance in GMES.

Among those Scenarios, Protected Areas Scenario focus the harmonisation works on schema mapping and transformation of the structure and geometry of datasets of Protected Areas, for the purposes of planning, management and tourism promotion, and demonstrates HUMBOLDT capabilities for geodata harmonisation in protected areas domain.

The HUMBOLDT project shows challenges both to geosciences research, covering topics in data harmonisation at a continental scale, and to economic and political management of such a large and heterogeneously composed consortium of partners. Nonetheless, the more relevant the challenges to face, the better the benefits which will surge from their solutions: benefits for specialised and non specialised users of spatial data, for policy-makers, planners and managers, for European citizens and their organisations, at a level which varies from local to regional to European are to be delivered through HUMBOLDT, during the year to come and even after the formal end of the project, fixed in early 2011.

References

- [1] Annoni, A., Smits, P.C. 2003. Main Problems in Building European Environmental Spatial Data. *International Journal of Remote Sensing*, vol. 24, no. 20, pp.3887-3902
- [2] Bernard, L., Kanellopoulos, I., Annoni, A. & Smits, P.C. 2005. The European Geoportal – One Step Towards the Establishment of a European Spatial Data Infrastructure. *Computers, Environment and Urban Systems*, vol. 29 (2005), pp. 15-31
- [3] Smits, P.C., Friis-Christensen, A. 2007. Resource Discovery in a European Spatial Data Infrastructure. *IEEE Transactions on Knowledge and Data Engineering*, vol. 19, no. 1, pp. 85-95
- [4] De Vries, M., Giger, C., Loidold, M. 2007. HUMBOLDT – State of the Art in Data Harmonisation and Data Management. Deliverable A3.5D1 of HUMBOLDT project (March 2007)
- [5] Portele, C., Van Oosterom, P., Bayers, E. et al. 2007. INSPIRE Methodology for the development of data specifications; *INSPIRE Drafting Team "Data Specifications"* (May 2007)

- [6] Villa, P., Reitz, T., Gomasasca, M.A., 2007. HUMBOLDT project: implementing a framework for geo-spatial data harmonization and moving towards an ESDI. *Geoinformation in Europe, Proceedings of the 27th EARSeL Symposium*, Millpress, pp. 29-36
- [7] Villa, P., Gomasasca, M.A., Reitz, T., 2008. HUMBOLDT project for data harmonization in the framework of GMES and ESDI: Introduction and early achievements. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Volume XXXVII, Part B4, pp. 1741-1746
- [8] Reitz, T., Holweg, D., Ludlow, D. et al. 2006. HUMBOLDT – Development of a framework for data harmonisation and service integration – Description of Work; Annex I of the HUMBOLDT project contract (October 2006)