



The HUMBOLDT project

towards the harmonisation of spatial information in Europe

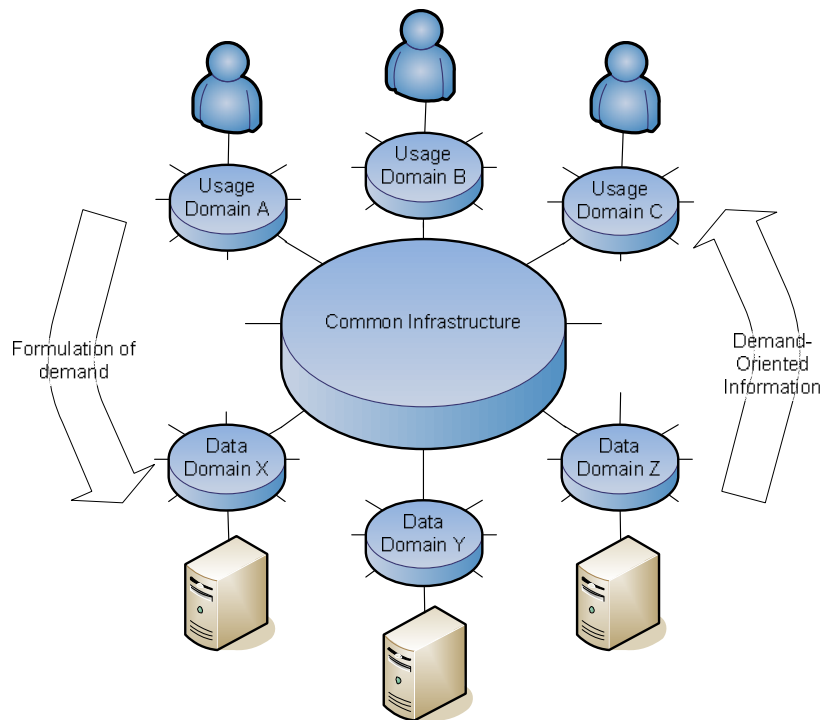
Dániel Kristóf, *Institute of Geodesy, Cartography and Remote Sensing (FÖMI), Hungary*

Eva Klien and Thorsten Reitz, *Fraunhofer Institute for Computer Graphics (FHG-IGD), Germany*

Christine Giger, *Swiss Federal Institute of Technology (ETHZ), Switzerland*

Santiago Cáceres, *ETRA Research and Development, Spain*

The HUMBOLDT Project (IP in FP6, 10/2006 to 09/2010)



▣ (Some of the) Objectives

- Transsectoral, Translingual, Transborder usage of geoinformation
- Common infrastructure to support transformation needs
- Support for INSPIRE (tools for data harmonisation) and GMES (theme-specific services in scenarios)

▣ Users' perspective on the SDI

- From a data-centric view to a usage centric view

HUMBOLDT – Facts

| | |
|-----------------------|--|
| Full title | <i>Development of a framework for data harmonisation and service integration</i> |
| Term | 48 months 01/10/2006 – 30/09/2010 |
| Effort | ~ 13.5 mill. € (~ 50% self-funded) ~ 110 person years |
| Contracting authority | Commission of the European Community FP6 – Aeronautics and Space (GMES) |
| Consortium | 28 partners (coordinator Fraunhofer IGD) |



HUMBOLDT Partners



Fraunhofer-IGD (GER), ETRA (ES), HSRS (CZ), Logica CMG (UK), IGN (F), Intergraph (CZ), ETH Zürich (CH), TU Delft (NL), Uni Rome (I), FÖMI (H), MARIS (NL), KTC (Lit), INI-GraphicsNet (GER), IFREMER (F), NERC/BODC (UK), HCMR (G), SMHI (S), UWE (UK), Telespazio (I), GISIG (I), TUM (GER), CNR-IREA (I), FMI (CZ), IGP (P), CLS (F), HiG (S), Promiteas (Cyp), Intergraph (GER)

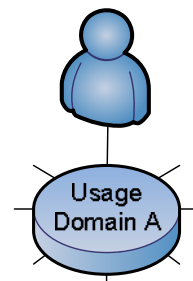
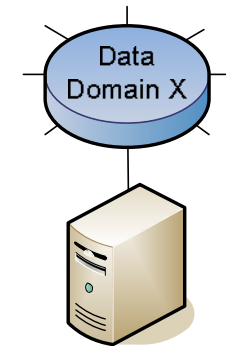
HUMBOLDT – Vision



- ▣ Format
- ▣ Spatial
- ▣ Graphical
- ▣ Object
- ▣ Language
- ▣ Object semantics
- ▣ Expected quality

“creating the possibility to **combine data from heterogeneous sources into integrated, consistent and unambiguous information products**, in a way that is of **no concern to the end-user**” (A 3.5-D1)

- ▣ Provision of information in the language area / semantics of the



about the

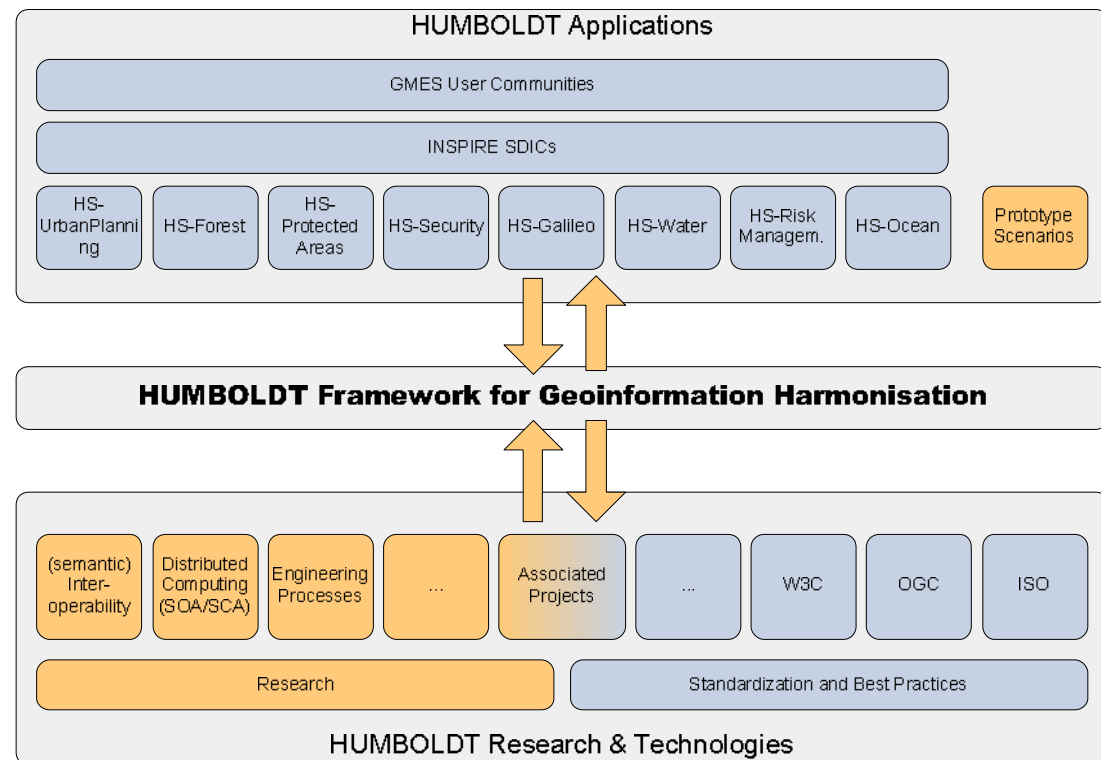
HUMBOLDT Implementation

Application-driven Approach

- Identification of user requirements
- Proof of concept in different domains (scenarios)
- Evaluation of framework against user requirements

Technology-driven Approach

- Technical concept, implementation, and evaluation of the framework
- Contributions and use of standards (OGC, W3C)

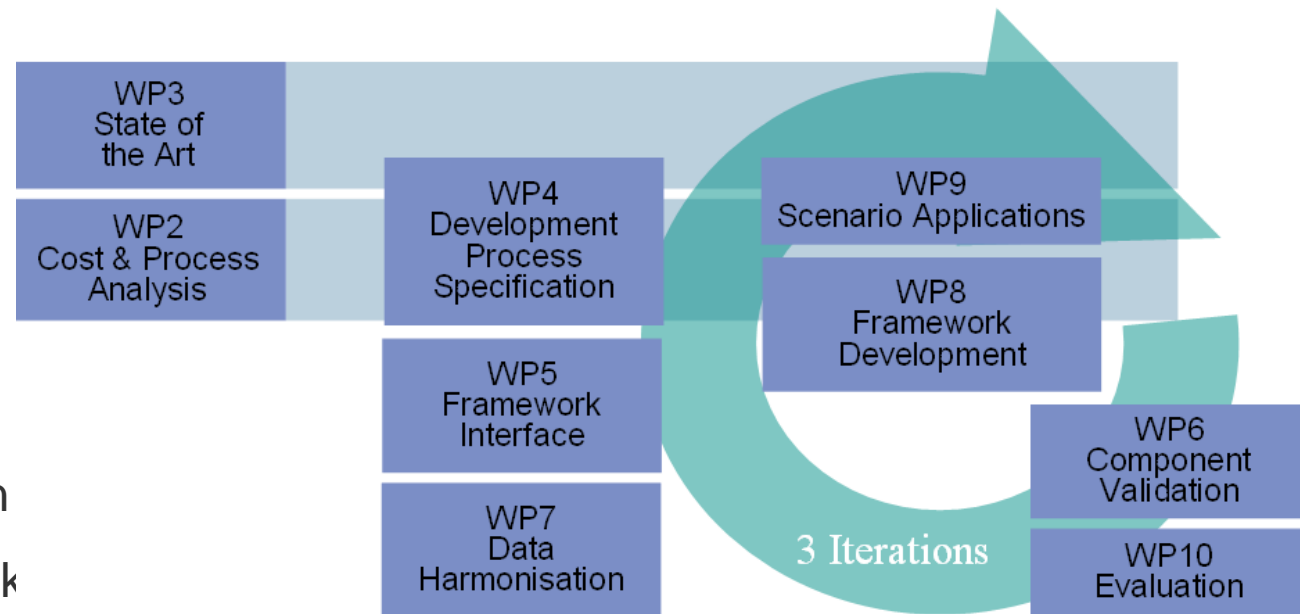


Technical Approach: Harmonisation Framework

Collection of tools for harmonisation (and use) of spatial data with a level of automation as high as possible

- Concepts
- Components
- Tools

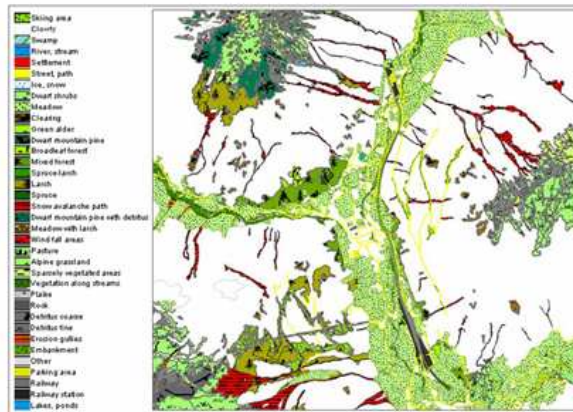
Provision as Open Source Framework



User-driven Approach: Scenarios



Tackling the Alpine Convention – Multi-Scale Monitoring of the Alps



List of Scenarios

- Border Security ERiskA
- Urban Planning Water Catchment
- Forest Ocean
- Protected Areas Galileo

Applies to each scenario :

- Implementation of an end user application using harmonised data
- Evaluation against user requirements
- Demonstrator for the framework

Some details about the scenarios

- ▣ **BorderSecurity** – revolves around supporting border police on the outer border of Europe, the focus is on the use of highly sensitive data, real-time sensor input and high availability of geodata from trans-border sources.
- ▣ **Urban Planning** – addresses environmental, architectural, and social needs of stakeholders in urban planning by offering them integrated access and visualization of such different data as sociodemographic information and city models, which makes this Scenario a perfect test for data integration from different application domains. Recently, Urban Atlas-related activities were also incorporated.
- ▣ **Forest** – centres on the management of forests in Czech Republic both also between countries on both sides of the Czech-German border, metadata harmonisation, geometric and semantic harmonisation and different sources of geodata are the aims.
- ▣ **Protected Areas** – highly different data sets to be used together, from biodiversity, land cover information to data for planning, management and promotion of protected Areas. These geo-information flows are identified at all management levels (EU, national; regional and local).
- ▣ **ERiskA** – aims at developing a cross-border risk management application for floods in the Lake Constance region (Austria, Germany and Switzerland).
- ▣ **Transboundary Catchment** – applications to manage transboundary catchments and to evaluate effects of water consumption or contamination in the context of the Water Framework Directive using the HUMBOLDT framework.
- ▣ **Ocean** – the largest Scenario in HUMBOLDT, focuses on cross-border, multi-language data, with high amounts of satellite data playing an important role. Distributed processing and 4D visualizations are used for the simulation of oil spills and other disasters. Includes 3 sub-scenarios.
- ▣ **Galileo** – makes use of the new GNSS Galileo to collect and distribute atmospheric data. These are required from several domains, use cases concentrate on healthcare applications.

Who are our users?

User roles:

- **GIS developer:** People who are doing programming jobs in the environment of GI Systems
- **GI data custodian:** people/institutions who are offering data which have been adapted to given standards (harmonised) – because of legal or market requirements
- **GI data integrator:** people who have to use heterogeneous geodata to meet the requirements of their daily job (e.g. integration of data for complex analysis)
- **End-user of geo data:** people who are working with already harmonised geodata or geodata that doesn't need harmonisation or integration at all
- **End-user of spatial information:** users at a laypersons level, e.g. people who are using online route planning services (not dealing with the real geo data)

■ Data harmonisation components

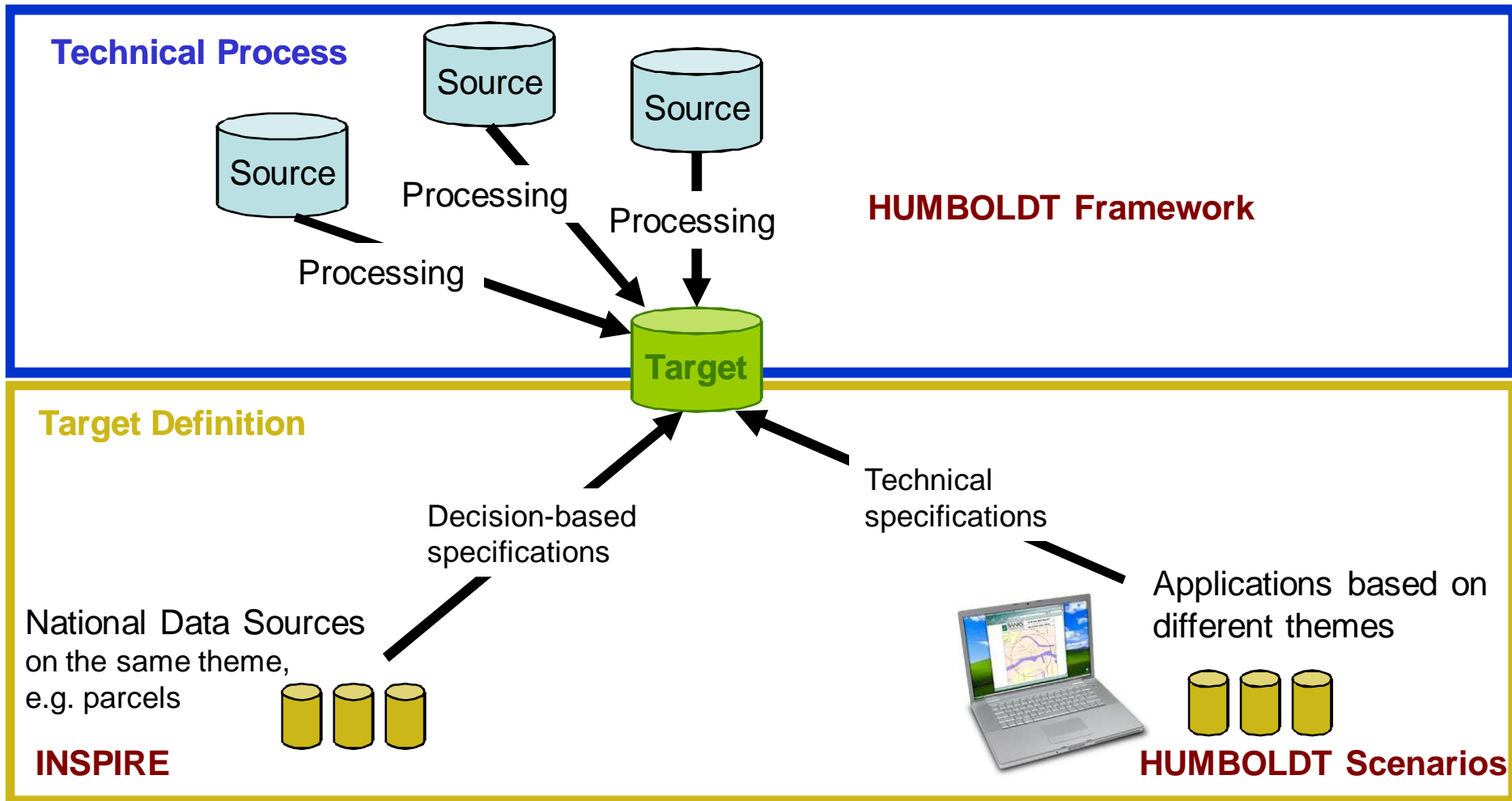
■ Data harmonisation components (INSPIRE, RISE)

| | | |
|--|----------------------------------|--|
| (A) INSPIRE Principles | (B) Terminology | (C) Reference model |
| (D) Rules for application Schemas and feature catalogues | (E) Spatial and temporal aspects | (F) Multi-lingual text and cultural adaptability |
| (G) Coordinate referencing and units model | (H) Object referencing modelling | (I) Data translation model/guidelines |
| (J) Portrayal model | (K) Identifier Management | (L) Registers and registries |
| (M) Metadata | (N) Maintenance | (O) Quality |
| (P) Data Transfer | (Q) Consistency between data | (R) Multiple representations |
| (S) Data capturing | (T) Conformance | |

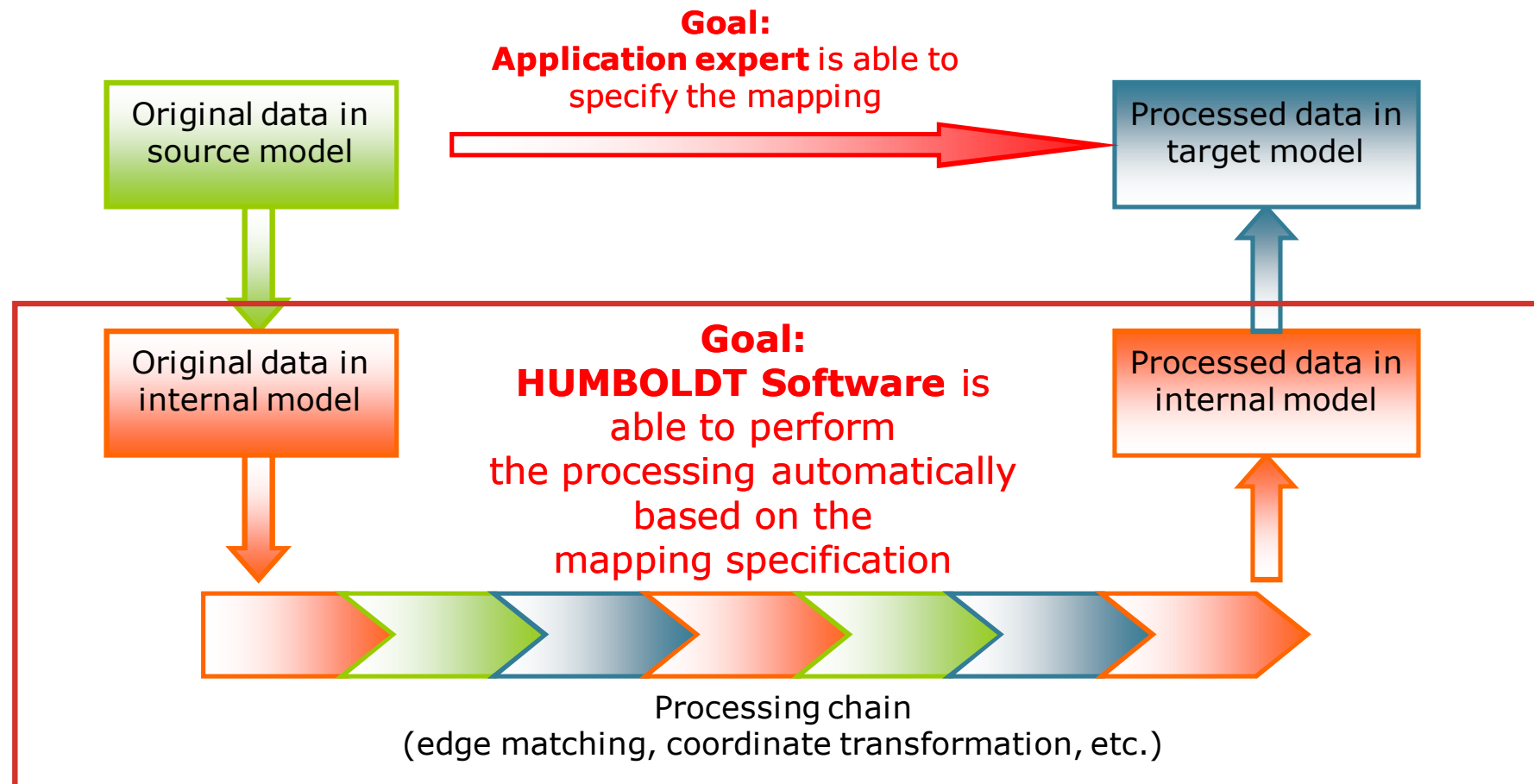
HUMBOLDT:

+ computational models (process models), their constraints and parameters
as 21st component

■ Data harmonisation processes



Implication for HUMBOLDT: Technical processing of data



HUMBOLDT products

- ▣ There are three levels of **software products** defined and implemented:
 1. **The HUMBOLDT Framework:** A set of software components that can be used by themselves or all together to create geodata harmonisation applications and that are re-useable in multiple toolset applications and scenario applications → **DEVELOPERS**
 2. **The HUMBOLDT Toolset:** A set of applications that support data modelling experts with all activities surrounding data harmonisation processes → **DATA CUSTODIANS, DATA INTEGRATORS, (end-users of geodata)**
 3. **The HUMBOLDT Scenario Applications:** Concrete End-User applications using the framework components and (adopted) versions of toolset applications → **END-USERS OF GEODATA / SPATIAL INFO.**

HUMBOLDT products (contd.)

▣ Further, **supporting products** are also important from the user perspective:

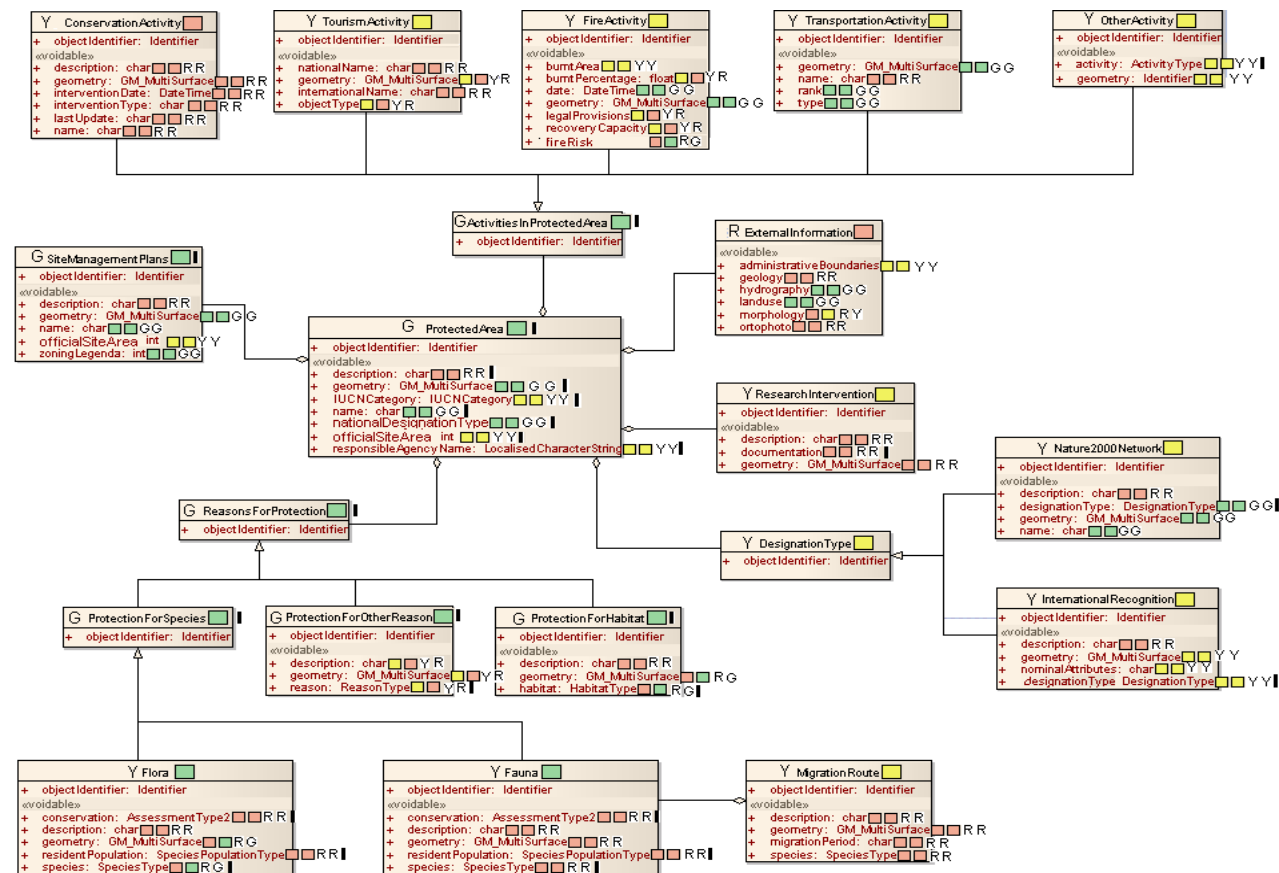
1. HUMBOLDT Documentation:

- Reports, deliverables and other documentary output
- Software documentation, „cookbooks”
- HUMBOLDT Wiki (currently used internally, can be made available to the wide public in the future)

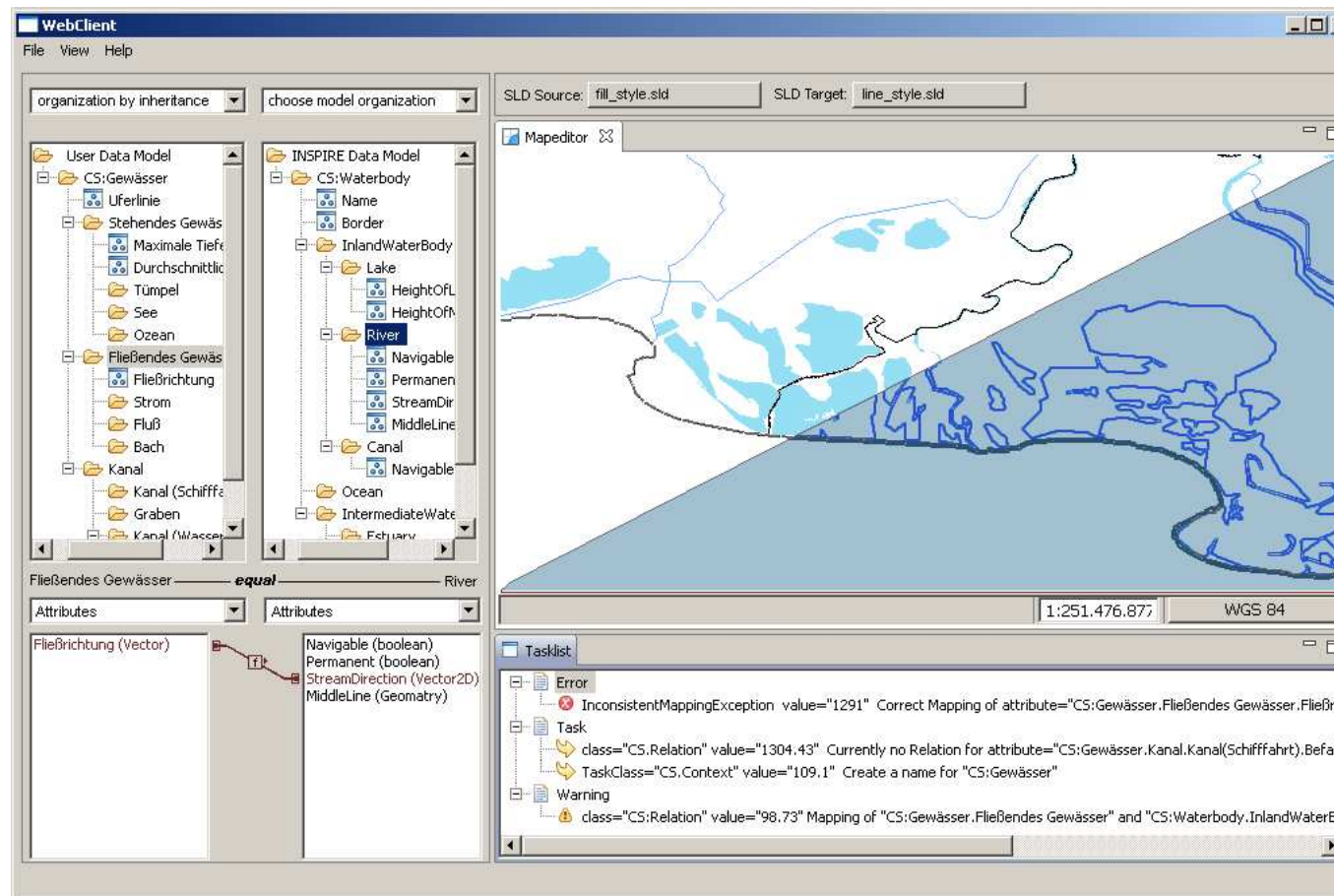
2. Training products:

- HUMBOLDT Training Platform
- Overall training on data harmonization, specific trainings on the HUMBOLDT project tailored to user needs

- **Toolkit: Schema definition and extraction** focused on the creation of INSPIRE-compliant, geographic application schemas
- **Application: HUMBOLDT GeoModel Editor**



- **Toolkit: Schema alignment** enabling a conceptual-level description of transformations needed to transform data from one application schema to another
- **Application: HUMBOLDT Alignment Editor (HALE)**

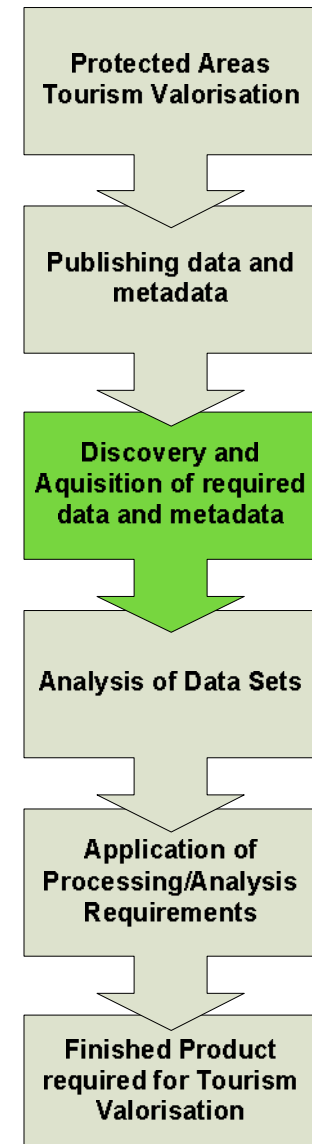


- **Framework: Automated discovery** of matching data sets to context – currently includes ~10 parameters that can be set, e.g Geometry, CTS, Keywords, ...)

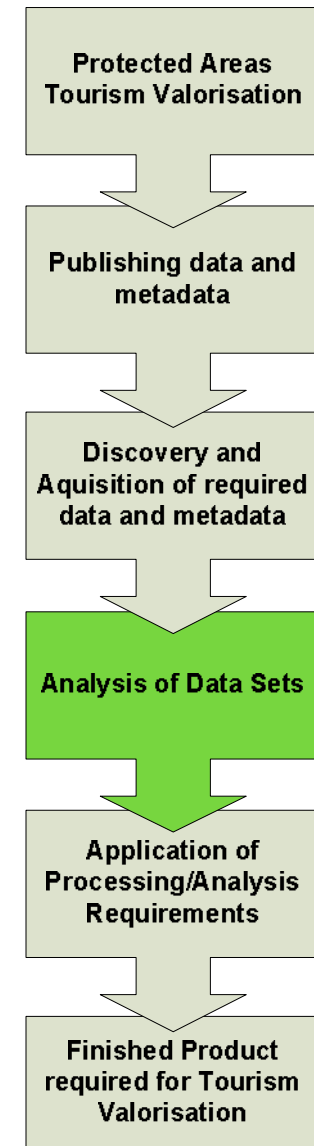
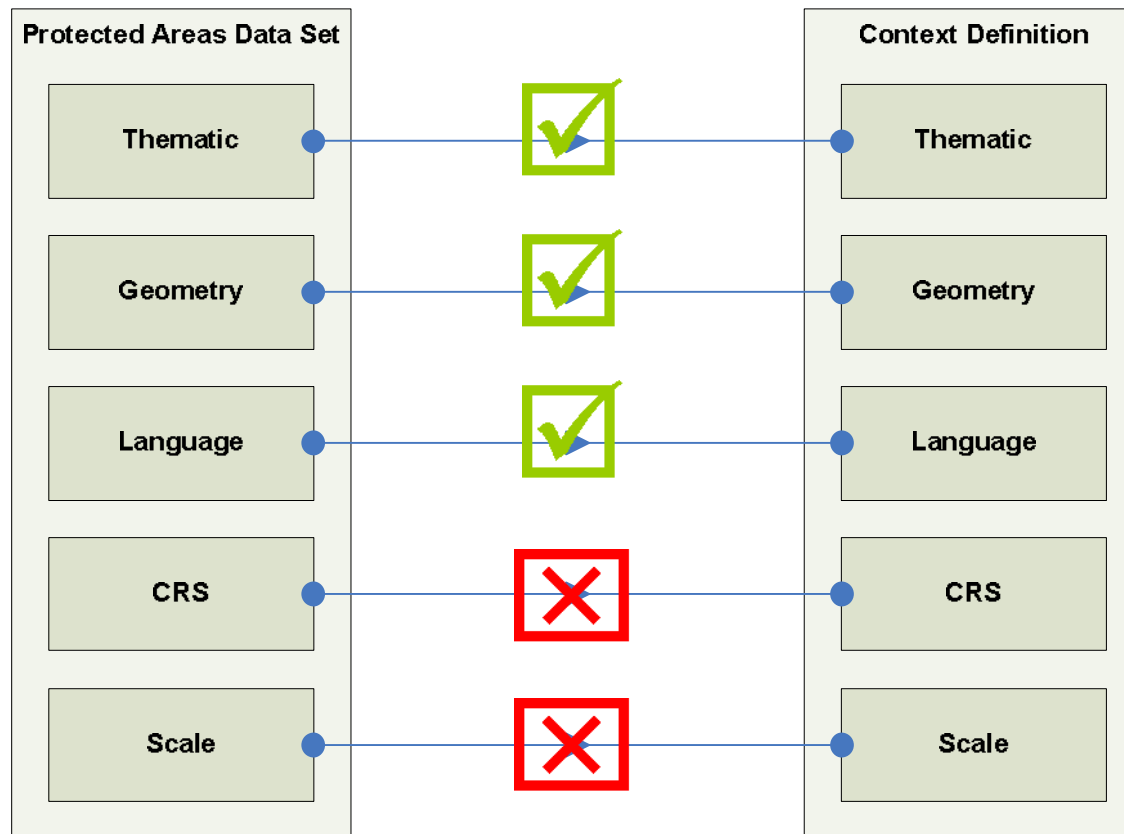
| A | B | C | D | E | F | G | H | I | J | K |
|--|----------------------------|--|-----------------------------------|--|----------------------|--|---------------|--|----------------------------------|-----------------------------|
| Information levels derived by TR on the basis of the proposed user story | Download | Services Names Proposed (code and filename) | Shapes/ies | Map preview server (openlayers) old layer name | Geometry | Projection | Data source | Dataset indicated in the system | restingVMS URL | ExistingMet a |
| 2 | Download | IT_Municipalities_20000_COMMU_2000.shp | | Direct link | Polygon | GAUSS BOAGA - MERKATORE TRANSVERSO COMPORWED GAUSS | Regional data | Administrative boundaries (BVerw NUTS level) sc:1000 | Direct link | Metadata |
| 3 | | IT_Municipalities_50000_LA_COMMU_5000.shp | | Direct link | Polygon | Datum INTERNAZIONALE (HA) PCFO) - ROMMAD | | | | |
| 4 | | IT_NUTS_2000 | PROVINCE_2000.shp | Direct link | Polygon | | | | | |
| 5 | | IT_NUTS_5000 | PROVINCE_5000.shp | Direct link | Polygon | | | | | |
| 6 | Download | IT_Municipalities_4_30_Administrative_Municipalities.shp | | | Polygon | Reference Datum: Transverse Gauss, EPSG:27462 VMS Projection: EPSG:27462, Datum: EDN, VFS Projection: EPSG:27462 | Local (part) | Dataset: Administrative boundaries (CA) PCFO) Scale: 1:25000 | VFS on direct link | Metadata |
| 7 | | IT_Planet_Map_2 | DT_CADP_Municipalities.shp | Direct link | | | | Dataset: Administrative boundaries (CA) PCFO) Scale: 1:25000 | Status: online VFS (GML version) | Direct link |
| 8 | Varying availability | | | | | | | DEM derived from the Geovector regional scale: 10000 | Fake VMS | |
| 9 | Download | IT_Municipalities_Park_2000 | DT_Park_Municipalities | | Point (with polygon) | Datum: Transverse Gauss, EPSG:27462 | Local (part) | Dataset: DTM, year: 2000 | Direct link | Metadata |
| 10 | Download link upon request | | | | Line | GAUSS BOAGA - MERKATORE TRANSVERSO COMPORWED GAUSS | Regional data | Elevation: 10000 | Direct link | Metadata |
| 11 | Varying availability | | | | | | | | | |
| 12 | Download | IT_Rivers_Regional_Catchment | ELEMENTI_ERORI.shp | Direct link | Line | GAUSS BOAGA - MERKATORE TRANSVERSO COMPORWED GAUSS | Regional data | Hydrography: sc:10000 | Direct link | Metadata |
| 13 | Download | IT_Rivers_Regional_Park_Catchment | IT_Rivers_Regional_Park_Catchment | Direct link | Line | GAUSS BOAGA - MERKATORE TRANSVERSO COMPORWED GAUSS | Local (part) | | | |
| 14 | Direct link | Chimney smoke | DT_Planet_Map_2 | | Line | Datum: Local, Height: Gauss | | Dataset: E. Hydrographic network Scale: 1:500000 | to data: VFS | To be provided |
| 15 | Download | IT_Regione_Liguria_Municipalities | zone_90.shp | Direct link | Polygon | GAUSS BOAGA - MERKATORE TRANSVERSO COMPORWED GAUSS | Regional data | Environment: Special Protection Area: sc:12000 | Direct link | Metadata |
| 16 | Request | | | | | | | Environment: Natura 2000 | Direct link | Metadata |



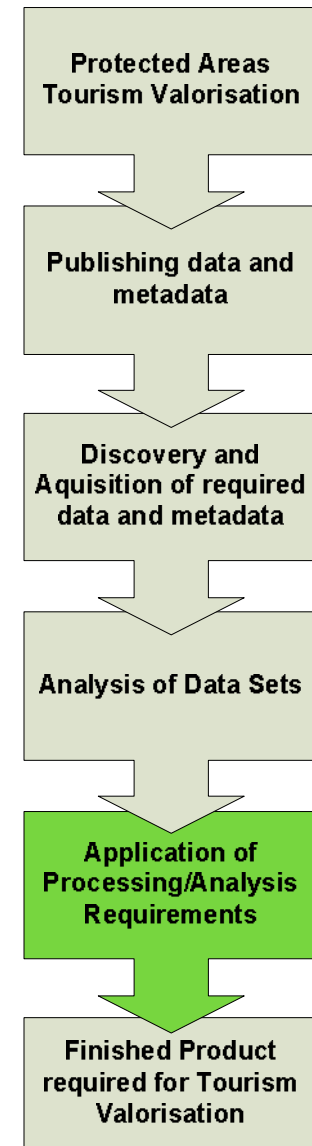
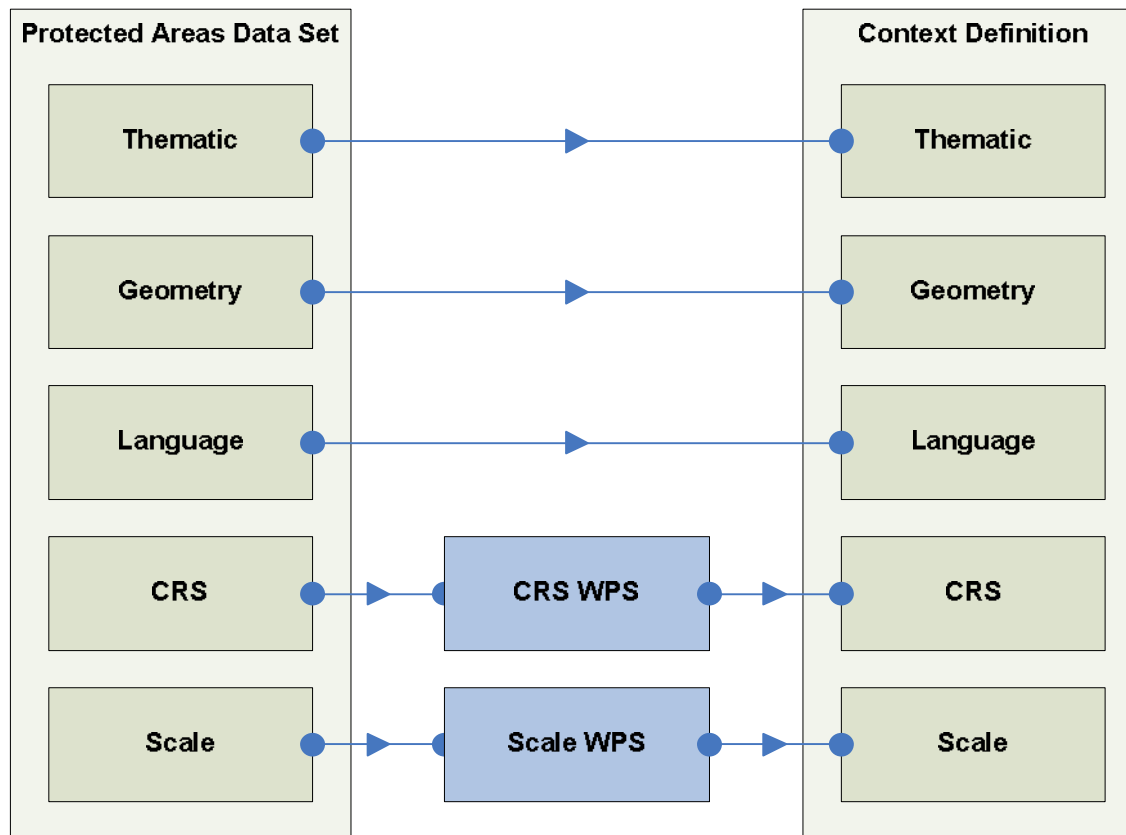
- Candidate Data Set I
- Candidate Data Set II
- Candidate Data Set III



- **Framework:Automated analysis** of discovered datasets for “fitness for use”
 - the system analyses through a set of rules whether dataset can be used directly or whether transformation is necessary.



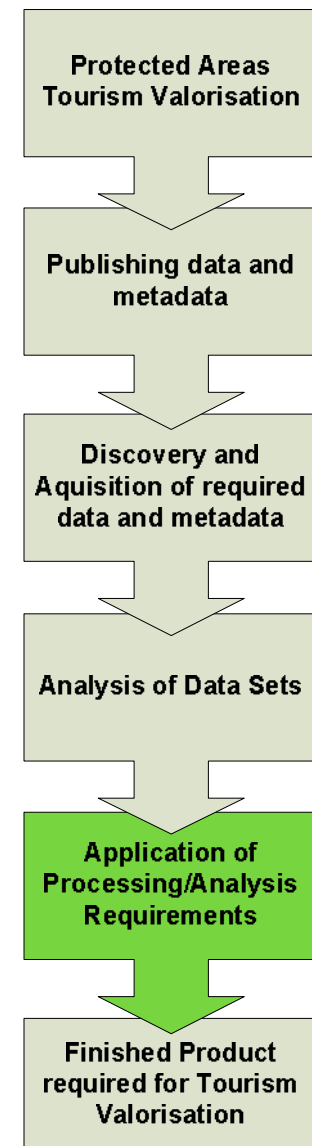
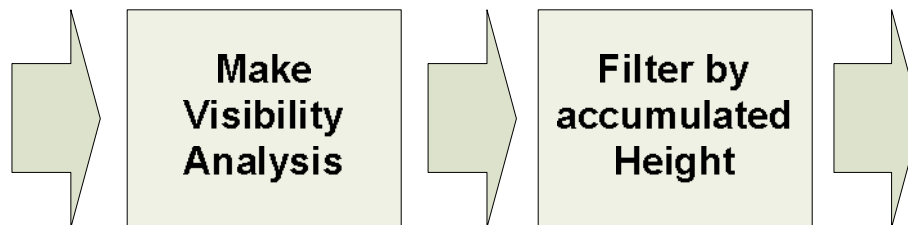
- **Framework: Automatic discovery and execution** of processing capabilities to resolve heterogeneities



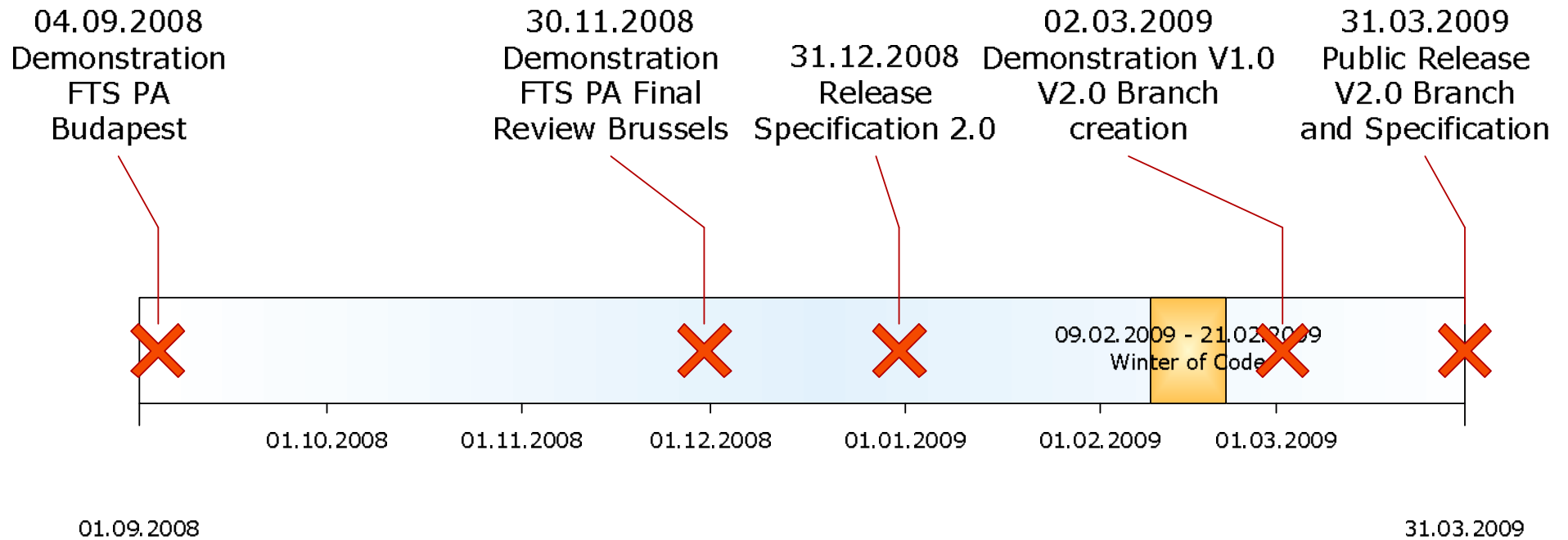
Framework: Application-Specific Processing

- Automatic execution of processing and analysis steps belonging to the domain methodology
- The ASP portion is **portable/data independent**
- Example from the Protected Areas scenario:

Mario: Filtering of routes by burnt areas visibility and accumulated height of a trail



Timeline for the FOSS release of the Framework



What is being released?

1. Everything under the V2.0 Branch of the Framework:

1. A „stable“ build based on the lowest version of everything under V2.0 (effectively V1.0 Final, as developed until beginning of March including the Winter of Code event)
2. A „development“ build representing the latest implementation efforts leading to full 2.0 implementation

2. Specification V2.0 and the supporting APIs and documents

3. Everything developed on that base.



Thank you for your attention!

Join our user community!

- ▣ Website: www.esdi-humboldt.eu
- ▣ Application Manager: kristof.daniel@fomi.hu
- ▣ Coordinator: eva.klien@igd.fraunhofer.de
- ▣ Project Office: po@esdi-humboldt.eu