



The HUMBOLDT Framework: FOSS and Standardization in a large-scale research & development project

CASCADOSS Symposium, Warsaw, Poland

16.06.2008

Thorsten Reitz

Fraunhofer Institute for Computer Graphics Research (IGD), Darmstadt, Germany

Overview

The HUMBOLDT Project

- What is the HUMBOLDT Framework?
- Who are the targeted users?

Standards and FOSS in HUMBOLDT

- What are the drivers in the project and it's environment for FOSS?
- Is FOSS a suitable basis for such an R & D Project?

The relation of Standardization and FOSS

- How do Standardization and FOSS influence each other?
 - Examples from Geo-related domains
 - Examples from elsewhere

Benefits and Challenges of using FOSS for geocentric Applications

- What are the next steps?

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

▣ (Some of the) Objectives

- Transsectoral, Translingual, Transborder usage of geoinformation
- Provision of common infrastructure components to glue data producer and consumer transformation needs together
- Support for INSPIRE (tools for data harmonisation) and GMES (theme-specific services in scenarios)

▣ User's perspective on the SDI

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

Key Elements of the Approach

Application-driven Approach

- Identification of user requirements
(incl. “communication skills”)
- Proof of concept in different domains

(scenarios)

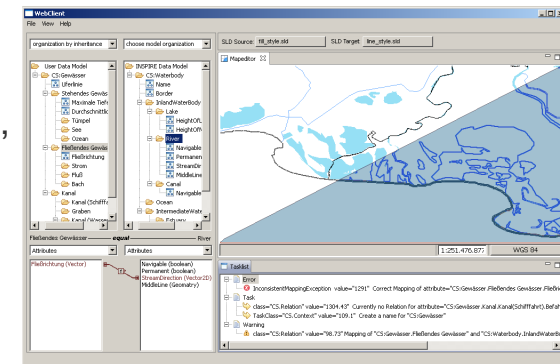
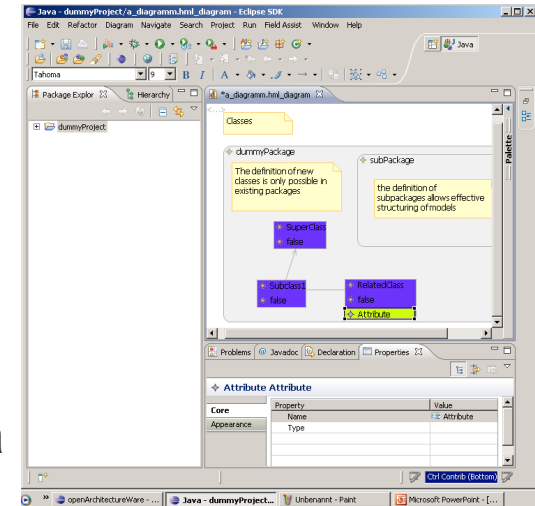
- Evaluation as to user requirements

Technology-driven Approach

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

What is the HUMBOLDT Framework?

- ▣ Concepts for supporting a Spatial Data Infrastructure in Europe, specifically for Service Integration and Data Harmonization
- ▣ Reference Implementation of these concepts as a collection of tools for harmonization of existing spatial data and services with a level of automation as high as possible
 - “framework for the technical process”: **Service Components** for the automatic transformation/translation of this geodata from one logical, physical and conceptual schema to one required for a task at hand
 - “supporting the target definition” in the scenarios: **Software Tools** for defining and exploring the relations between heterogeneous sets of geodata



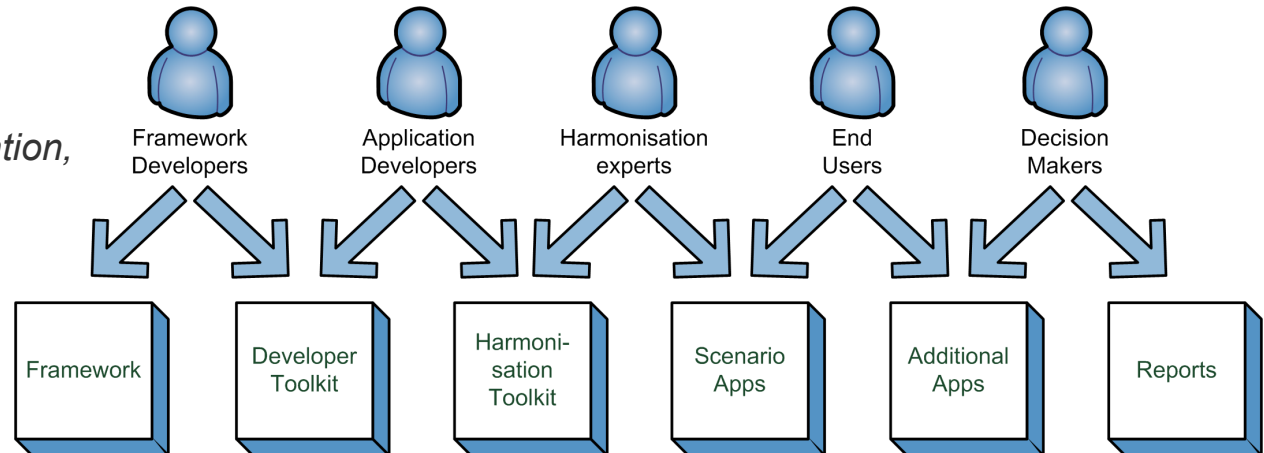
HUMBOLDT Users

Direct users

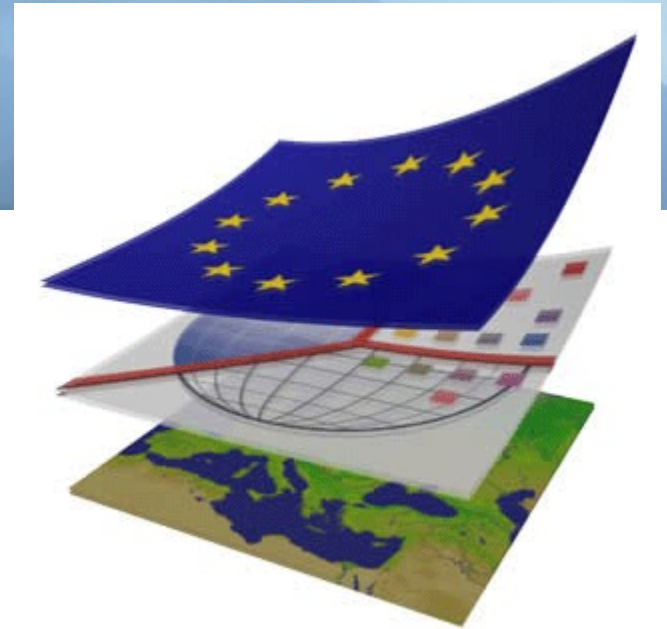
- ▣ *(Framework) Developers*: Contributors to the development of the HUMBOLDT Framework
- ▣ *Data Integrators/Application Developers*: Integration of data sets for applications and development of applications based on the HUMBOLDT Framework components
- ▣ *Data Custodians*: Perform data harmonisation for data in their organisation and are experts in at least one geoinformation domain
- ▣ *End Users of Geodata*: Experts from the geoinformation domains who perform analysis, modelling and visualisation based on harmonised geodata resources

Indirect users

- ▣ *End Users of Geoinformation, such as Decision Makers*: Use services developed by Application Developers



Standards and FOSS in HUMBOLDT



FOSS/Open Standards Drivers *in* the Project

▣ (End) Users:

- FOSS:
 - Easy access to software without long administrative processes
 - Sufficient quality of software
 - Sustainable access to software, even if producer goes out of business
 - Low cost of licenses
- Standards:
 - Usage of established software based on those standards
 - Exchangeability of software based on standards
 - Mandated in some Areas such as Public Administration

▣ Developers (especially Industry):

- FOSS:
 - Easy changeability of software
 - No licensing issues with shared IPRs
 - Low risk (no long-term contracts, up-front licensing costs)
- Standards:
 - Usage of established software

Key Drivers in the Environment: Public Administration Users

- ▣ Public administration represents the biggest portion **of currently registered HUMBOLDT users.**
- ▣ In Public Administration, a change on multiple levels is happening.
- ▣ Drivers of change in public administration:
 - Reduced Budgets and lay-off of personell
 - Structural change
 - Introduction of private sector-type controlling instruments
 - Development of e-Government
 - Expectations of citizens and enterprises
- ▣ Suitability of FOSS in this change
 - **Open standards for exchange & high sustainability** mandated
- ▣ => Public Administration requires FOSS and represents a core user group!

FOSS and open standards in HUMBOLDT

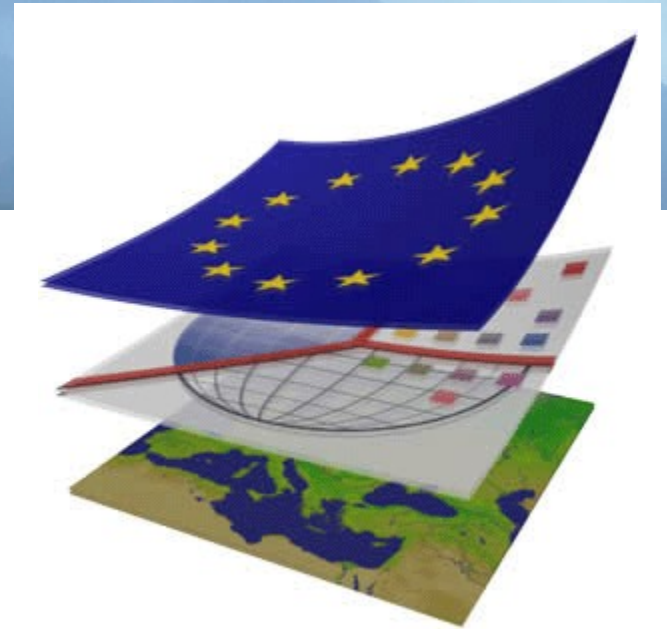
▣ Release of FOSS and standards by the project:

- Determination of licensing:
 - Release of the Specifications, Framework and Toolset under LGPL (ca. 01/2009)
 - Scenarios distributed with proprietary licenses
- Business Model:
 - varies with partner interests, network effects are core
- Standardization:
 - Active collaboration with OGC, INSPIRE, W3C

▣ Usage of FOSS and standards in the project:

- State of the Art Evaluation of commercial and FOSS tools and libraries
- Evaluation of standards for their importance and suitability
- **Result: “*Handbook of Standards*”: Summary of Input standards to HUMBOLDT**

The Relation of Standardization and FOSS



How do Standards influence the Project?

- ▣ Handbook of Standards: results of an evaluation of both tools and standards
- ▣ Selection of Standards for the Handbook:
 - Based on a relevance analysis within the project's scenarios
 - Based on a set of qualitative and quantitative criteria, such as:
 - Addressed HUMBOLDT requirements*
 - Maturity:** How often is the Standard used? Are these research projects or also Production usage?
 - External Evaluation:** Have others evaluated the standard? With what result?
 - Update Frequency:** How active is the community supporting the standard's continuous development?
 - Breaking changes:** How often are versions released that are incompatible with existing standards versions?
 - Longevity:** How old is the standard?
 - Support:** How is the standard supported in terms of consultancy and implementation?
 - Design Criteria:** Modularity, High cohesion on one topic, Extensibility
- ▣ The Handbook is updated every 9 months within the project.

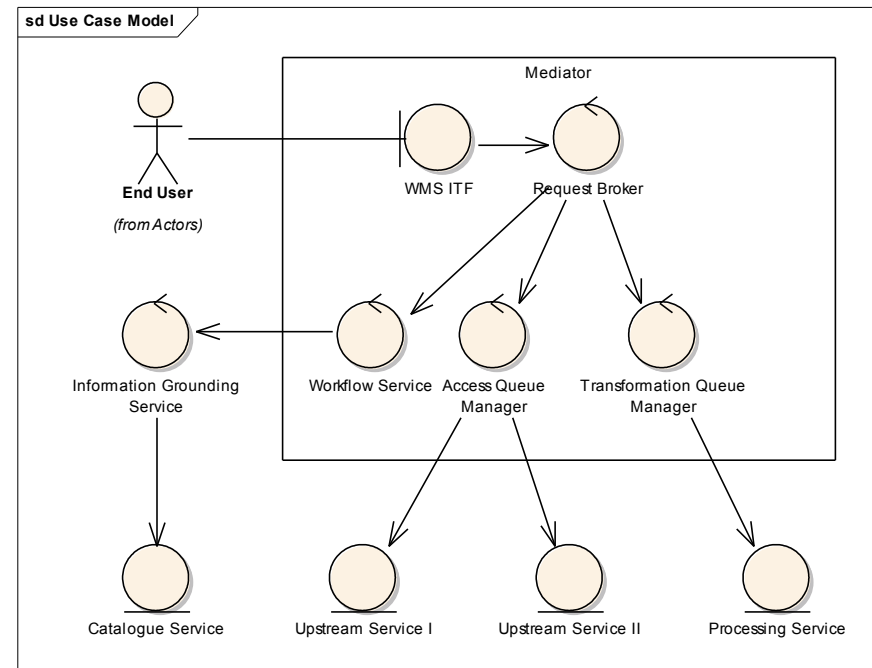
Which Standards actually influence the Project?

Geo-specific:

- Interface Specifications (CSW, WMS, WFS, **WPS**...)
- Data & Metadata formats (GML, ISO 19115 & ISO 19119...)
- Abstract Specifications (ISO 19103, ISO 19110, ISO 19113...)
- INSPIRE Implementing Rules...

Common:

- W3C/OASIS Web Service Standards (SCA, SOAP, WSDL, XML, XML Schema...)
- Java-centric (JPA, EJB3, JTA...)
- Process-Centric: ISO RM-ODP



What Impact do those Standards have?

General:

- Alignment of internal processing models to standardized models wherever possible to avoid expensive mismatches
- Usage of Interfaces as defined by the OGC where possible

Positive:

- Core standards are robust and useful and are directly used
- Communication in the project and its environment is made easier

Negative:

- Often, (especially the geo-centric) standards match only a subset of actual requirements
- Tool and Library support can be lacking which leads to increased effort for standards-conformity (example: GML)
- Some newer standards are not yet mature (such as WPS)
- Some (abstract) standards allow a lot of interpretation in implementation

How does the project influence Standards?

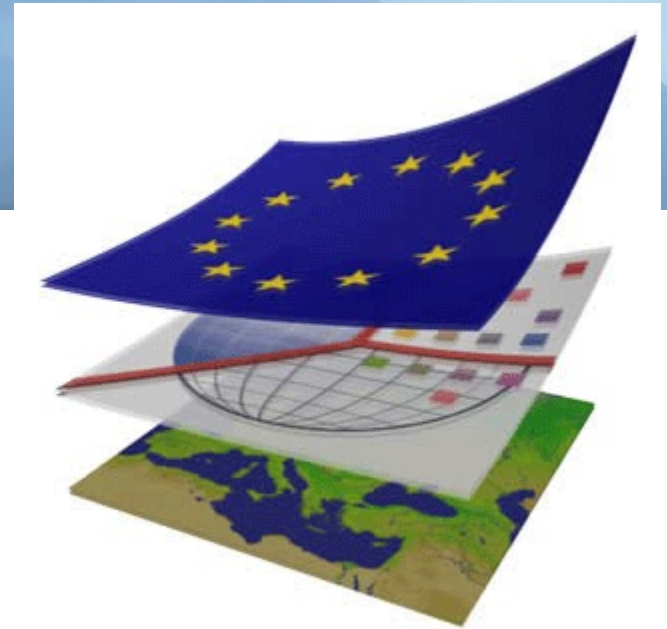
□ General:

- In FOSS, there has been a tendency to **not** follow standards but rather to suggest an alternative solution that is easier to implement and handles most use cases anyway
 - One reason is the focus on quick results and fast iterations instead of full standards compliancy in FOSS projects
 - Not-invented-here-syndrome still going strong, but sometimes integrate-or-make decisions are valid
 - Example: EJB vs Lightweight Containers
- Sometimes, these alternative approaches became best practice and then even standards
 - Example: Hibernate vs JDO
 - Example: Simple Features Profile of GML3.1

□ In HUMBOLDT: Evolution instead of Revolution

- In FOSS, GI standards are mostly accepted -> no major conflict as in app./web containers
- Work in OGC, specifically, SWE and WPS, investigating Geosemantics WG
- Contributions to W3C developments

Benefits, Challenges and next Steps



Summary of FOSS role and benefits

- ▣ FOSS important on „basic layer“: Operating System, Apache, ... and Containers, less so on top level (Applications Clients and Servers)
- ▣ FOSS is very important on meta-tools level: IDEs...
- ▣ FOSS is currently mainly used in common administrative processes, less so in domain-specific processes
- ▣ Without FOSS and Open Standards, a project like HUMBOLDT would not be feasible!

Next steps

- ▣ Improve collaboration between (java library) projects (OSGeo)
 - Better connection of Open Standards to Open Software

- ▣ Collaborative R&D within multiple concurrent research projects

- ▣ Build strong frameworks and libraries for geo-specific purposes

- ▣ Strengthen interoperability by supporting open formats (GML)
 - most important file format in FOSS still seems to be shape!

 Thank you for your attention!



Fraunhofer Institut
Graphische
Datenverarbeitung

Thorsten Reitz
M.Sc.

Department Geographic
Information Systems

Fraunhoferstraße 5
64283 Darmstadt
Telefon +49 (0) 6151/155-416
Fax +49 (0) 6151/155-444
thorsten.reitz@igd.fraunhofer.de
www.igd.fraunhofer.de/igd-a5