



Inspire Helsinki 2019

#inspirehelsinki2019

A framework for Building Solutions on INSPIRE

Thorsten Reitz (wetransform), Giordano Colombi
(AELER), Christopher Hönn (wetransform)

A brief intro to the challenge

“By the beautiful blue Danube”

Giordano Colombi, AELER



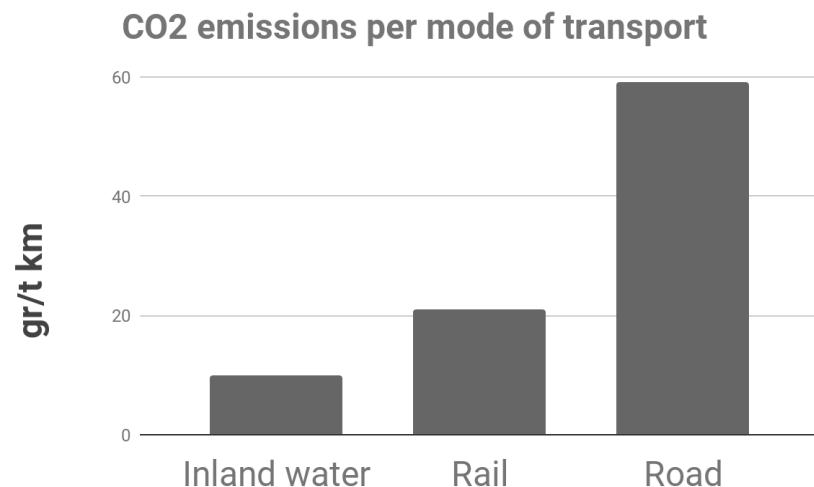
AELER

The Beautiful Blue Danube

In today's **supply chain logistics**, there is more and more public and political pressure to decrease the **carbon footprint of freight transport**.

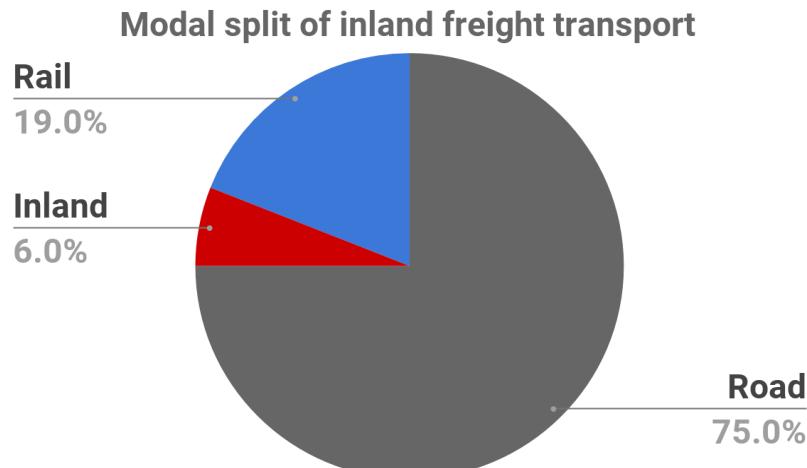
Fact 1

Maritime shipping is the world's most carbon-efficient form of transporting goods.



Fact 2

Inland water transport accounts for 6% of freight transport.



Sources:

<http://www.worldshipping.org/industry-issues/environment/air-emissions/carbon-emissions>
https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Freight_transport_statistics_-_modal_split

The Beautiful Blue Danube – Our roadmap

Q: How can we invert this trend? Would it be feasible to build an Intemodal Routing Model to prove the efficiency of inland waterways?

39.3 million tons of goods were carried on the Danube and its tributaries in the year 2018

Explore datasets related to transport infrastructures with focus around the Danube and Rhine

External factor might affect the goods transportation (e.g. climate)

Integrate them with other relevant geospatial datasets from EU countries

Think of an appropriate cost function for the problem to be minimized

Time, fuel cost, shipping cost?

Build the **Intermodal Routing Software** that shows the efficiency of rails and waterways usage

The main advantages of the inland waterways transport are: reliability, low environmental impact, high potential for decreasing the number of accidents, traffic jams, noise and air pollution

Sources: <http://www.plovput.rs/file/danube-stream/common-danube-report-2018.pdf>
<http://journals.univ-danubius.ro/index.php/research/article/view/3469/3921>

Data Usability - how INSPIRE could be helpful

INSPIRE Geoportal

Central European access point to the data provided by EU Member States and several EFTA countries under the INSPIRE Directive

- > 32k data sources
- > 25k downloadable data sources
- organized in **INSPIRE themes**



Q How can various datasets structures be harmonized to be used for intermodal routing?

Q How can INSPIRE Geoportal be used for addressing the intermodal routing problem?

Q What data are missing? Which kind of data would you like to have to better handle this thematic?



Our difficulties



Complicated to get access to the data



Legal restrictions on the usage



Poor documentation

Our use case

Got 54 result for the word
“Danube” but...

The screenshot shows the INSPIRE GEOPORTAL Resource Browser interface. At the top, there's a logo for the European Commission and the text "INSPIRE GEOPORTAL Enhancing access to European spatial data". Below this, the URL "EUROPEAN COMMISSION > INSPIRE > INSPIRE GEOPORTAL > Resource Browser" is visible, along with links for "Change log for this tool" and "Change log for validation reports and validation logic".
Filter: Options include "all resources" (selected), "errors", "warnings, no errors", "no issue", "Only metadata resources" (unchecked), and "Hide metadata resources" (unchecked).
Search: A search bar contains the text "Danube".
Add selection Criteria: A dropdown menu is open under "Resource Type", showing "dataset(54)", "series(1)", and "service(3)".
Current selection criteria (after applying the filter): "(x) Danube"
Pagination: "1 to 10 of 58 - 10 per page".
Results: Two datasets are listed:

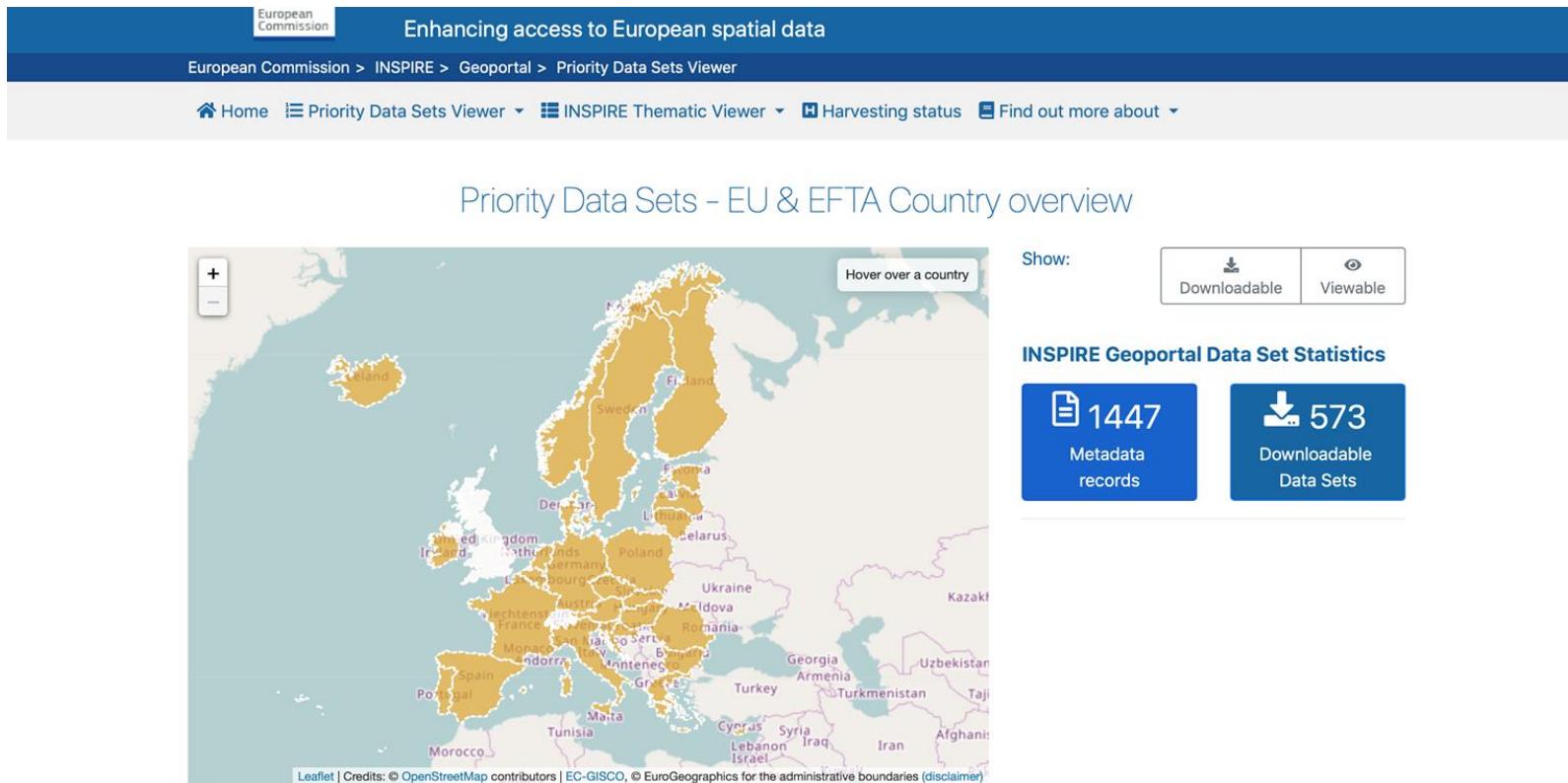
- (dataset) Terrain Model of Danube Delta Biodiversity Reserve 2011** (National Institute for Research and Development — Danube Delta — Tulcea)
Download Options | Inspire Metadata | Original Metadata | Evaluation Report
This data set is in the form of a grid in the Danube Delta Biosphere Reserve territory.
- (dataset) Electronic navigation maps (The River Danube River R.A., AFDJ Galari R.A.)**

... you have to **check manually** the **quality** and the **quantity** of the data for all the datasets

In order to get access to the data you need to reach out to several parties.
Very time consuming.

Our use case

In order to build an **optimization model** that works well, gathering the same kind of data for all the countries involved in our process was a crucial issue for us.



- Let's assume we would like to know in which country there are dataset related with the world "street"
- There is no way to filter the map using this criteria
- The mechanism didn't fit the needs of our use case.

Finding and Accessing Data

Thorsten Reitz, wetransform

Christopher Hönn, wetransform

Data providers challenges vs. Data user challenges

Data Requirements



- Published data and metadata
- Harmonized and interoperable data
- Relevant data for defined uses cases
- Presented in a way that reflects the use case (and not the data characteristics)

Our Experience finding data for the challenge

Observations:

- Initial responsiveness is slow (11 work days median)
- Only about 40% of contacts answered
- There are no standard license models or fees for usage of data as part of a cloud app or within a competition
- Average time to resolve usage rights for the challenge was 9 weeks with 6 to 14 contacts



Potential Solutions:

- Use standardized licenses and transparent conditions
- Have internal process to deal with licensing requests effectively

TASK 1: Can you help us find and access the data?

- ? Which data sets can you identify/find to solve this challenge?
- ? What is their license? Can we use them in our application?
What will be the costs?
- ? How can we access them?
- ? Whom do we need to contact?

Making Data Usable: Benefits and Challenges

Thorsten Reitz, wetransform

Better Documentation

...helps making data usable.

Observations:

- More than 50% of data sets we harmonise do not have any written documentation
- If there is documentation, it is mostly written for an internal audience in a limited set of languages
- There are only few experts on the data available, even in concrete projects

Enhancements through harmonisation :

- Documented concepts in multiple languages
- Conceptual documentation, e.g. UML models
- Detailed technical specifications available

The screenshot shows the INSPIRE Interactive Data Specifications interface. At the top, there's a logo for 'INSPIRE - Infrastructure for Spatial Information in Europe' and the title 'D2.8.III.2 Data Specification on Building – Draft Guidelines'. Below this, there's a detailed metadata table with columns for Title, Creator, Date, Subject, Publisher, Type, Description, and Contributor. Further down, there's a section for 'Favorite objects details' listing 'Building' and 'Buildings' with their respective descriptions and application schemas ('Building 3D' and 'Building 2D'). At the bottom, there's a footer with links to various INSPIRE services like Web Site, Legislation, Library, Forum, and Thematic clusters, as well as news and events, tools, and data specifications.

Consistent Quality Assurance

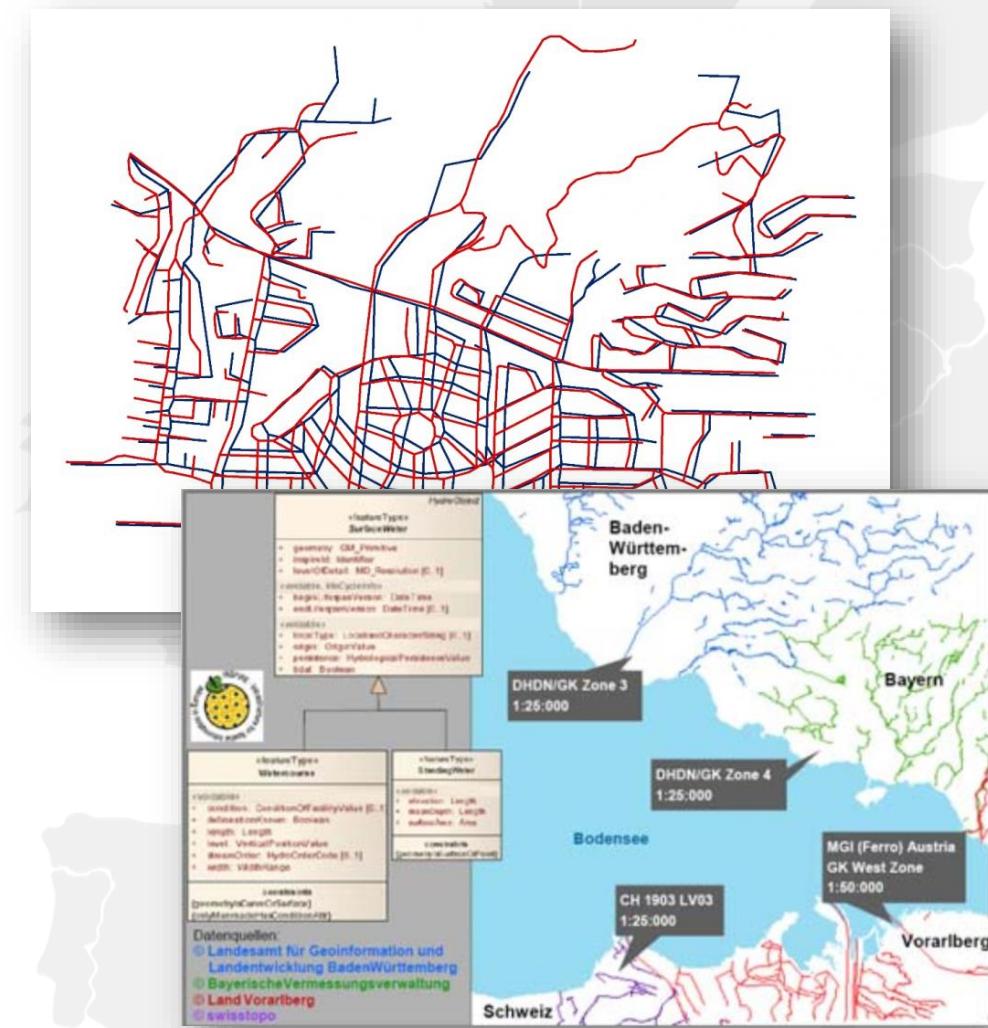
...makes data sets from multiple sources usable.

Observations:

- Most data sets have had only limited quality assurance and have errors in positional accuracy CRS, consistency, completeness and topology
- Quality standards are not documented anywhere or have never been defined
- Implicit standards (e.g. relational database schema) are the most common form of QA

Enhancements through harmonisation:

- CRS issues are identified and resolved
- Internal and external Topology issues are resolved
- Consistency and completeness are improved



Clearly defined semantics

...enable re-use of solutions and analytic processes.

Observations:

- In about 60% of data sets, semantics are defined in a local legal context and represent that well
- In 40% of data sets, semantics are not well defined, or there are mismatches to definitions

Enhancements through Harmonisation:

- Usage of standard Code Lists/Taxonomies
- Data can be combined with other data
- Tools and processes can be re-used
- High-quality, low-mismatch standard-to-standard harmonisation possible

The image shows two screenshots of the INSPIRE Registry. The top screenshot displays the 'Access Restriction' entry, detailing its ID, version, label, definition, governance level, status, themes, application schema, extensibility, and other formats. The bottom screenshot shows a taxonomy tool interface with a sidebar for 'Code list values' and a main panel for 'Taxonomy' showing asserted and inferred terms related to 'LocatorDesignatorTypeValue'.

Minerva Intelligence ACE

...but it doesn't come for free

Criterion	State	Factor
Do we know all data sources?	Yes	1,0
Do we have documentation of each data source that we can understand?	No	1,4
Is the data complete to deliver the target theme?	Yes	1,0
Is data quality sufficient?	Yes	1,0
Is there already a mapping table or other conceptual alignment documentation?	No	1,4
What format are the data sources in?	Shapefile	1,0
How do we access the data sources?	File	1,0
What is the number of mappable types in the source?	10	2,5
What is the number of mappable types in the target?	4	1,0
Average number of mappable properties	10...50	1,2
Does the target model use a topology?	No	0,7
What is the source data volume?	1GB...10GB	1,4
Final Multiplier		5,9
Base Project Price	12,0 h	1.500,00 €
This Project Indication	70,6 h	8.820,00 €

Easier Harmonisation

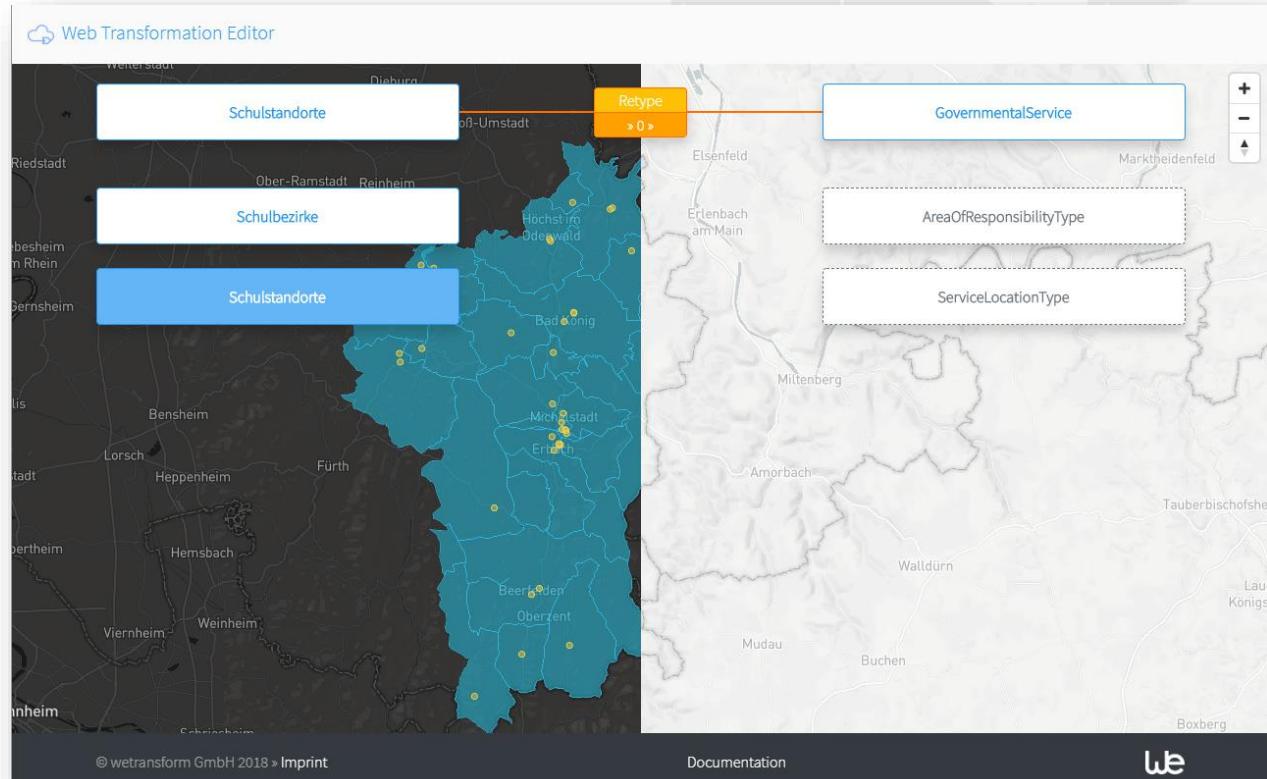
... lowers the entrance hurdle for data providers.

Observations:

- Many organisations don't harmonise their data (perceived difficulty, limited direct usefulness)
- Gaps in specifications decrease (semantic) interoperability

Potential Solutions:

- More & Better Code Lists/Taxonomies
- Fix issues in data specifications and ETS faster
- Share/Re-Use harmonisation projects
- Increase adoption of regional standards
- Increase automation and data-driven optimization



TASK 2: How can you help AELER get harmonised data?

- ? Do you have the resources (capacity and capability) to perform harmonisation?
- ? Which technical issues do you have with the tools you are using?
- ? Are there technical or semantical issues with the data specifications that block you?
- ? How do you prioritize data harmonisation?

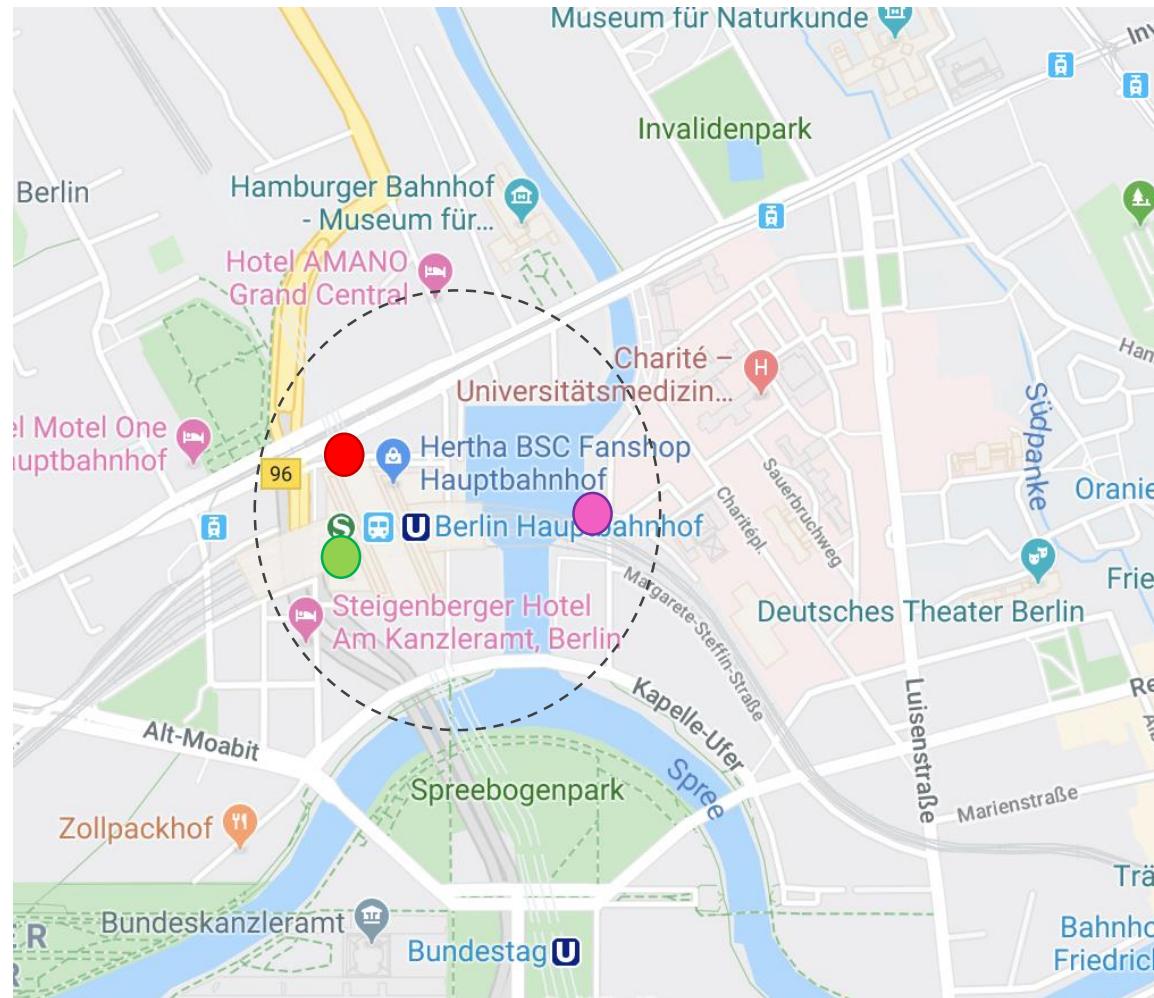
Making Data Useful: AELER Requirements

Giordano Colombi, AELER

Working with Geospatial Data – an Example

The three bullets
red, **green** and **pink**
represent the
Berlin Central Station
coordinates in 3
different data Sources

Unfortunately, The
semantic is not
harmonized so there is
not a unique way to
identify such Station.



To be sure that the 3 points
belong to the same station
we need to use some data
analytics techniques

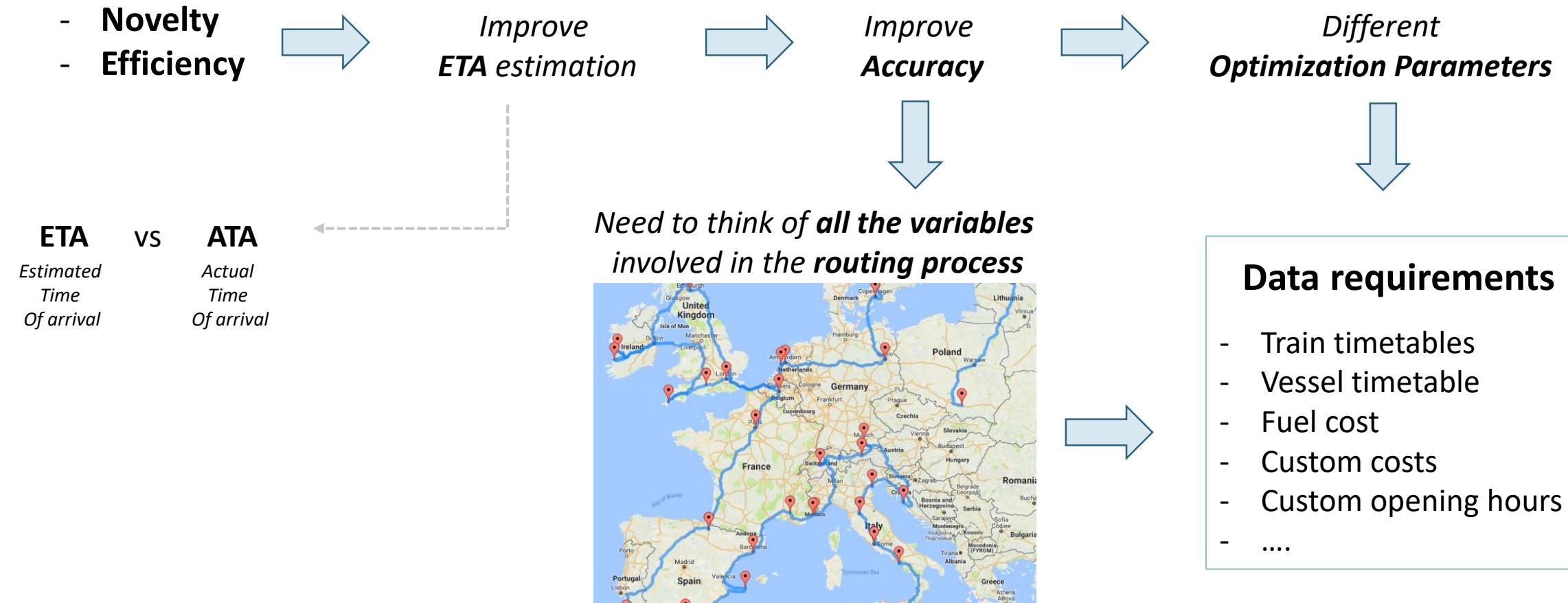
- **Geolocation**
- **NLP**

In both cases we have a
threshold problem.
**High chance to commit
errors**

We need a fast and smart
way to quickly identify
which point belongs to
which train station

AELER Requirements – Intermodal Routing

- Our Scope build an Intermodal Routing software
- Why? incentivize the usage of alternative modes of transport in Europe
- How?



TASK 3: What can you do to make the data useful?

- ? Would you go beyond minimum INSPIRE requirements to host specific (smart/adaptive) APIs?
- ? Would you also provide derived data products? Under what conditions?
- ? In what kind of applications would you like to see your data being used?

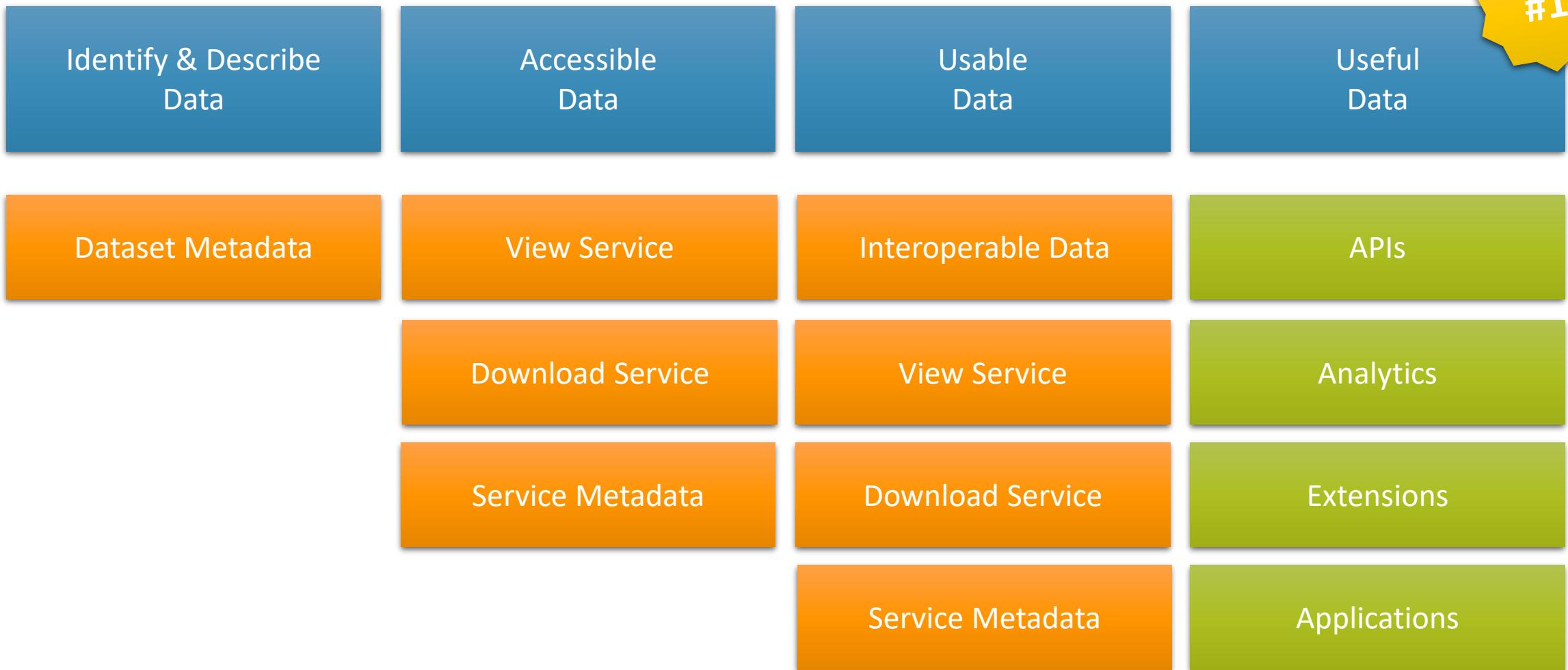
Summary: Building Solutions on INSPIRE

Thorsten Reitz, wetransform

Maturity of SDIs as Data Platforms

SDI Dimension		accessible	usable	useful
Specification		<ul style="list-style-type: none">International: INSPIRENational: GDI-DERegional: GDIS	<ul style="list-style-type: none">International: INSPIRENational: GDI-DERegional: GDIS	<ul style="list-style-type: none">Defined use cases/stakeholdersOptimized content/deliveryPrioritisation
Implementation		<ul style="list-style-type: none">MetadataView & Download ServicesRealtime APIs (STA)	<ul style="list-style-type: none">Interoperable Data setsValidated Data sets	<ul style="list-style-type: none">ExtensionsObject-centric UX (Linked Maps)Analytic APIsAI/BI Data Optimization
Maturity	Annex I + II National Level			
	Annex III National Level			
	Annex I, II, III State Level			
	Annex I, II, III Municipal Level			

INSPIRE Building Blocks



... and their implementation: Metadata

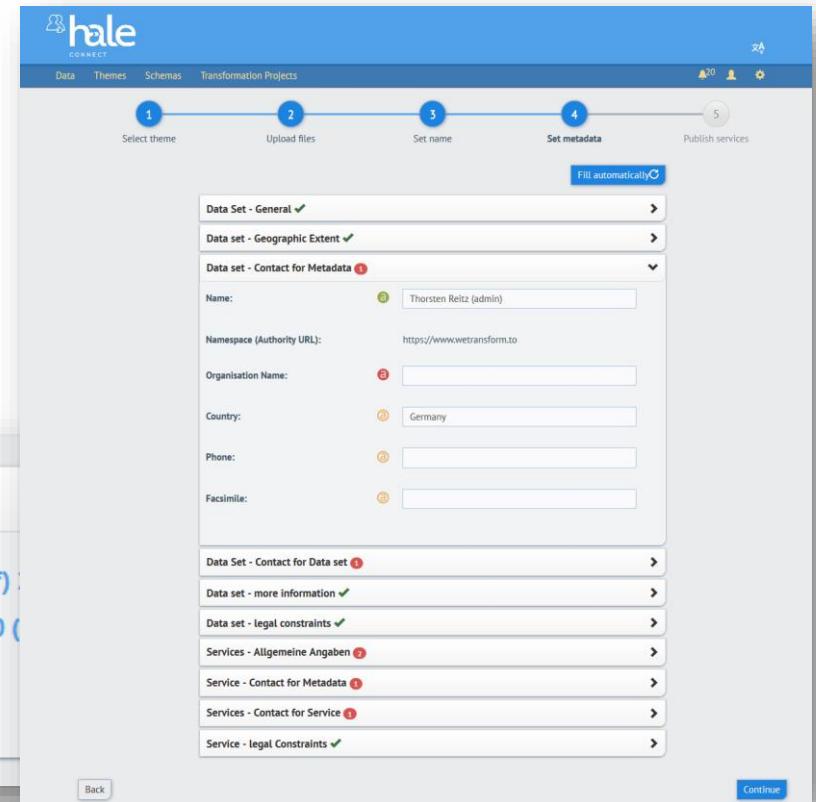
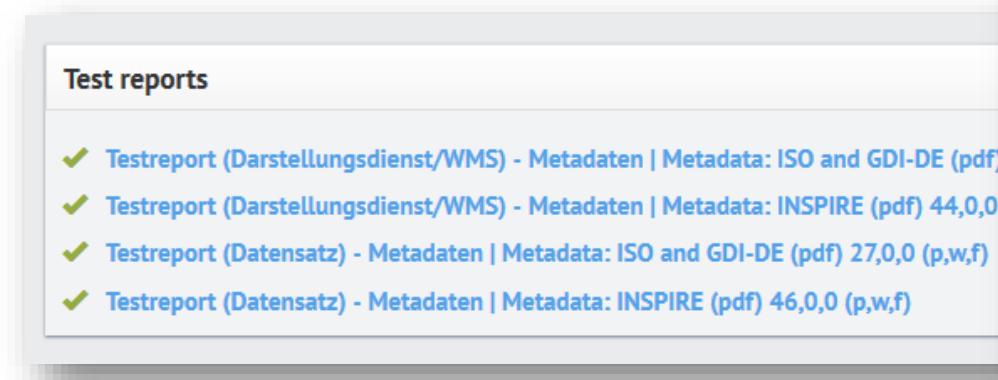
Identify & Describe
Data

Accessible
Data

Usable
Data

Useful
Data

- Automated Generation & Validation
- Re-use where available
- Consistent to multiple profiles (ISO, INSPIRE, ...)
- Continuous updates
- Automated publishing to Geoportals/ catalogues



Network Services

Identify & Describe
Data

Accessible
Data

Usable
Data

Useful
Data

- Automated Publishing
- One Dataset, many APIs
 - WFS, OGC Feature API, Atom
 - Attachments
 - WMS, WCS, STA, ...
- Extended Capabilities
- Service Coupling
- Scalable
- Monitoring



INSPIRE Interoperable Data

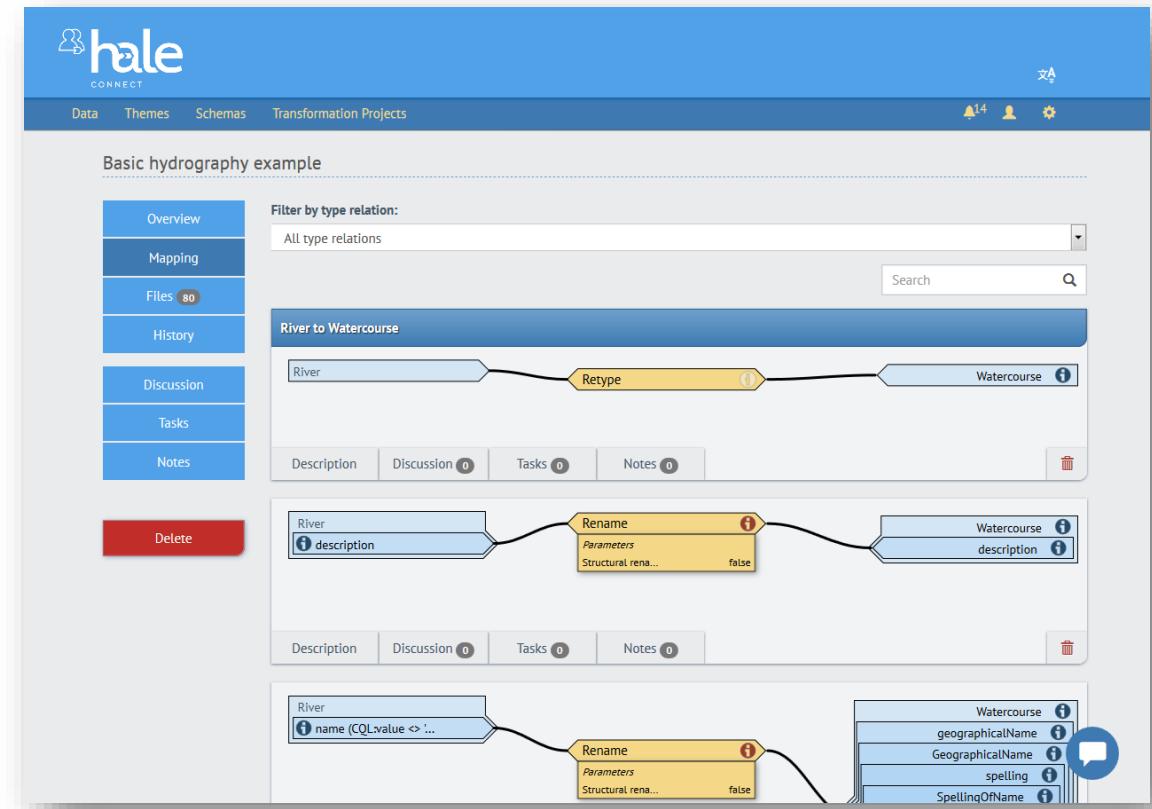
Identify & Describe
Data

Accessible
Data

Usable
Data

Useful
Data

- Full harmonisation
 - Geometry
 - Semantics
- Validation
 - Schema
 - ETF (Consistency, Constraints)



Beyond compliance to Usefulness

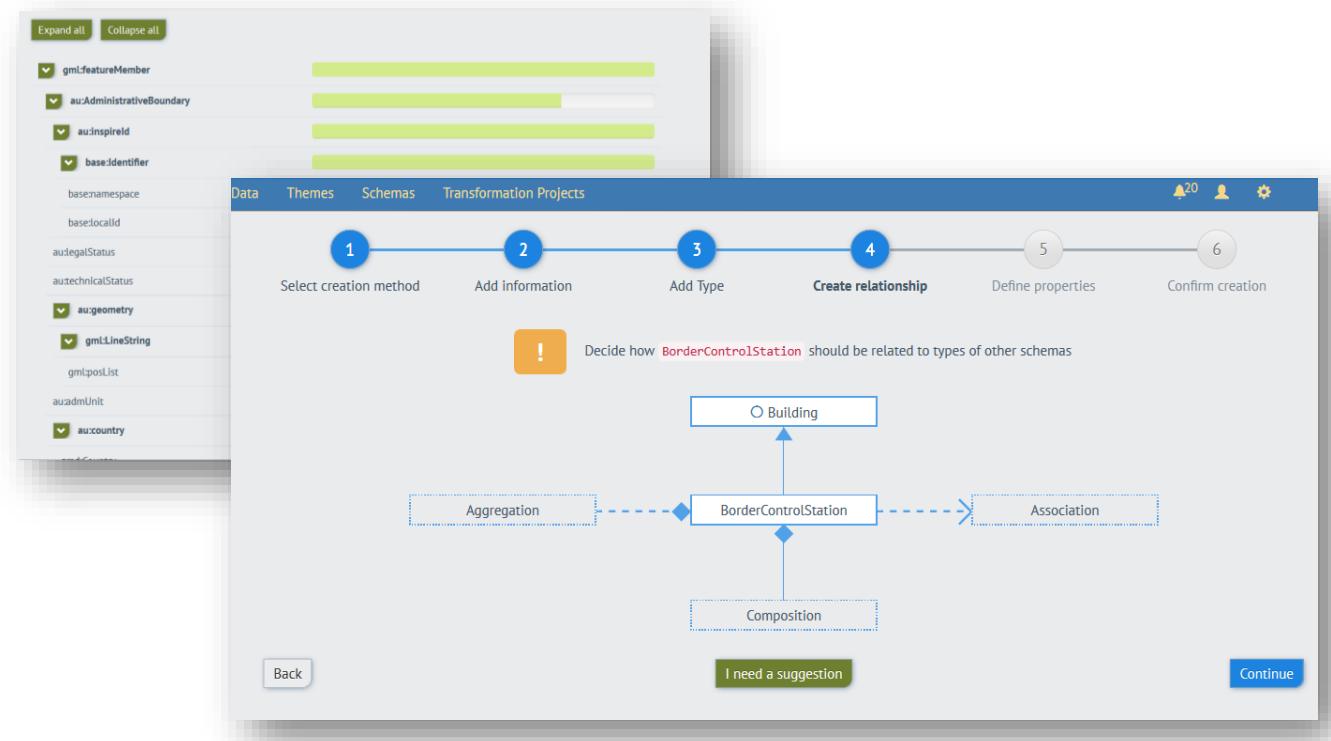
Identify & Describe Data

Accessible Data

Usable Data

Useful Data

- Optimize Data for AI/Analytics
- Provide Data via additional APIs
- Extend data models to match business processes
- Go beyond the Map to make the most of the data



... in one integrated workflow

The screenshot illustrates a five-step integrated workflow for dataset creation:

- Select theme**: Shows basic dataset metadata like Author (Joint Research Center (JRC)), Type (GMI Application Schema), Version (4.0), and creation date (13.04.2017). A sidebar displays file statistics.
- Select data source**: Shows a placeholder for dataset source selection.
- Set name**: Shows a placeholder for dataset naming.
- Set metadata**:
 - Dataset metadata**: A mandatory field "Enter URL" with an orange warning icon and a note: "Please enter the URL to the metadata that should be used for this dataset".
 - Service metadata**: Options to "Fill automatically" or view service contact details.
- Set execution rules**: Shows a preview of the dataset's INSPIRE components and their values.

Dataset metadata (Step 4):

Please enter the URL to the metadata that should be used for this dataset.

Enter URL

Service metadata (Step 4):

Fill automatically

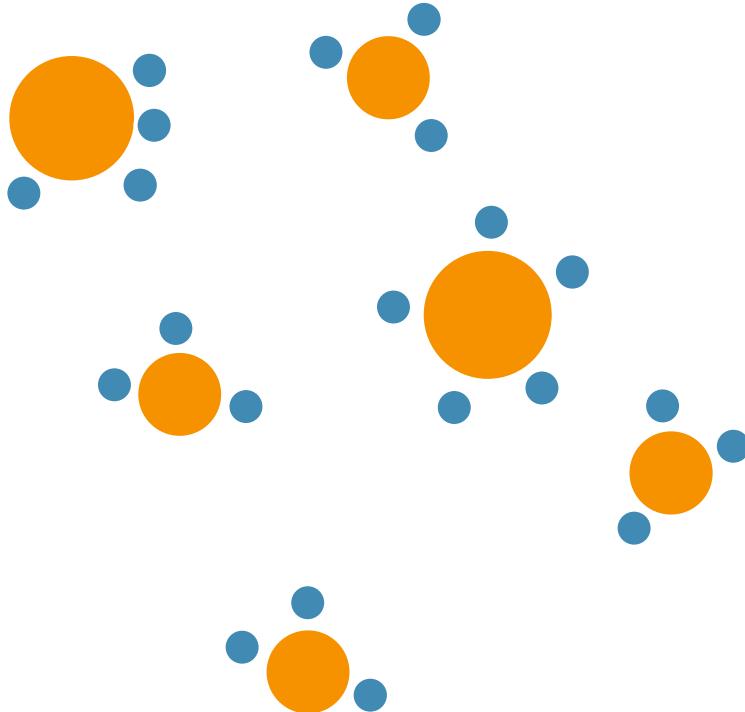
- Services - Allgemeine Angaben ✓
- Service - Contact for Metadata ✓
- Services - Contact for Service ✓
- Service - legal Constraints ✓

INSPIRE components (Step 5):

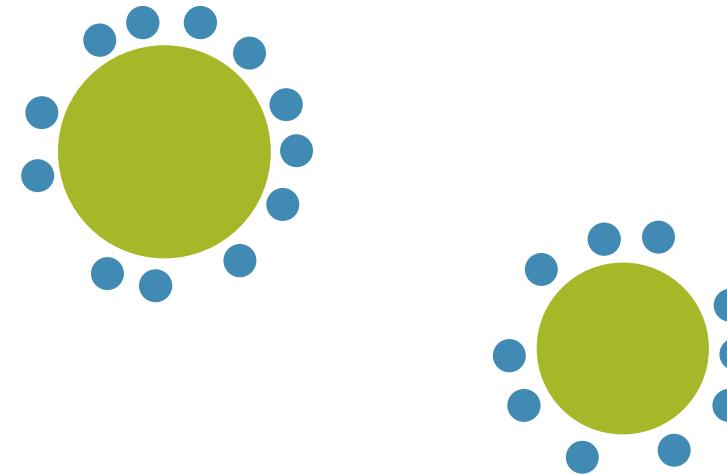
	Address
component (5)	1 Settlement_10086353 2 AdminInitName_1 3 ThoroughfareName_16178144 4 PostalDescriptor_21430382 5 AdminInitName_11026621
inspireId	localId: 11271634 namespace: http://www.eu.gov.si/INSPIRE/ad4.0
locator	designator: designator: 15 designator: type: address_number level: site_level
position	default: true geometry pos: 15.320875 46.262286 method: by_administrator specification: building

Usable, Standardised Data → Standardised Processes

Today (Fragmented):



In the future (Platform on Platform):



- + Effective Tool Development
- + Faster distribution of process enhancements
- + Lower Costs

Enable Solutions

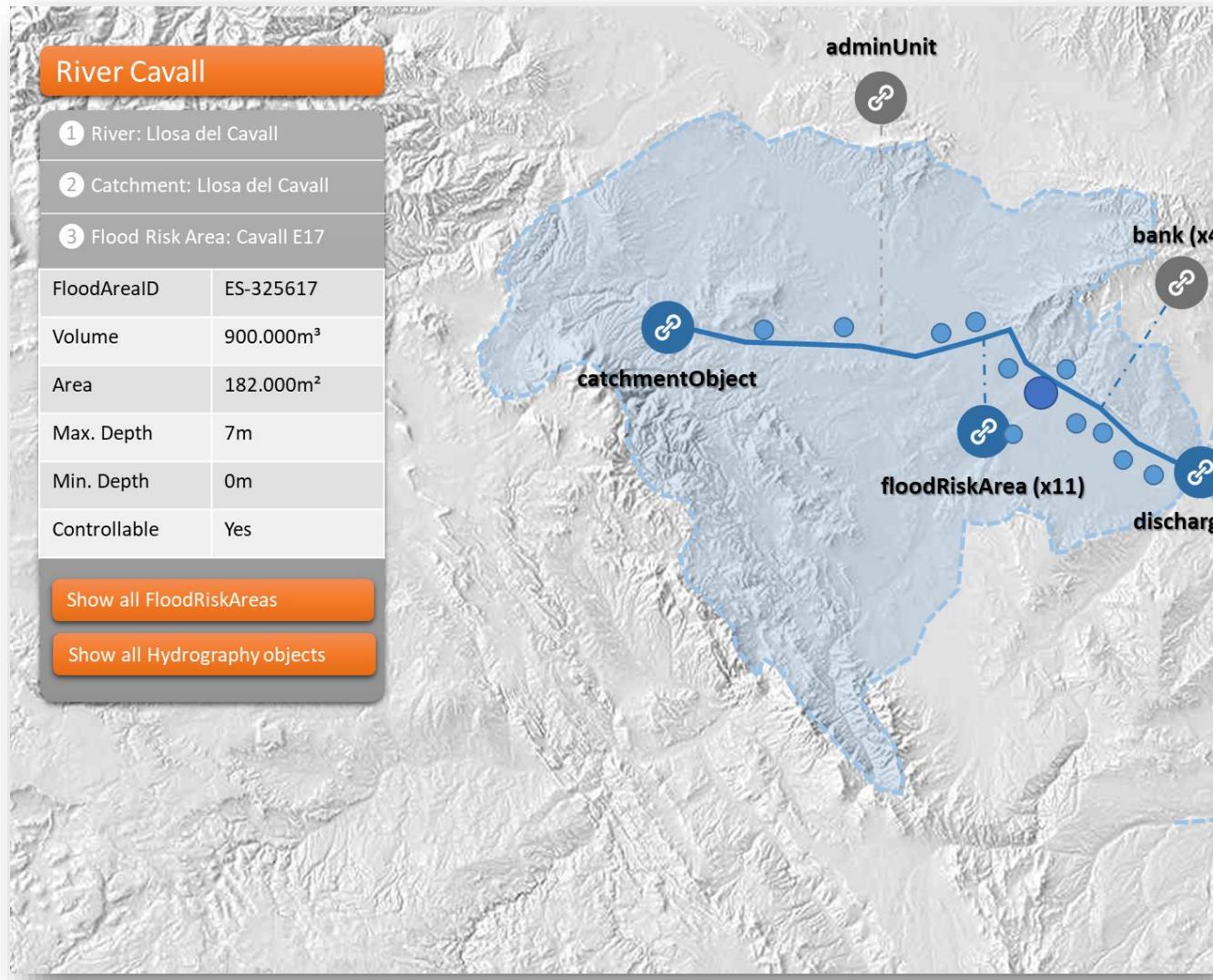
...to build an ecosystem of INSPIRE-based apps and professional tools

Observations:

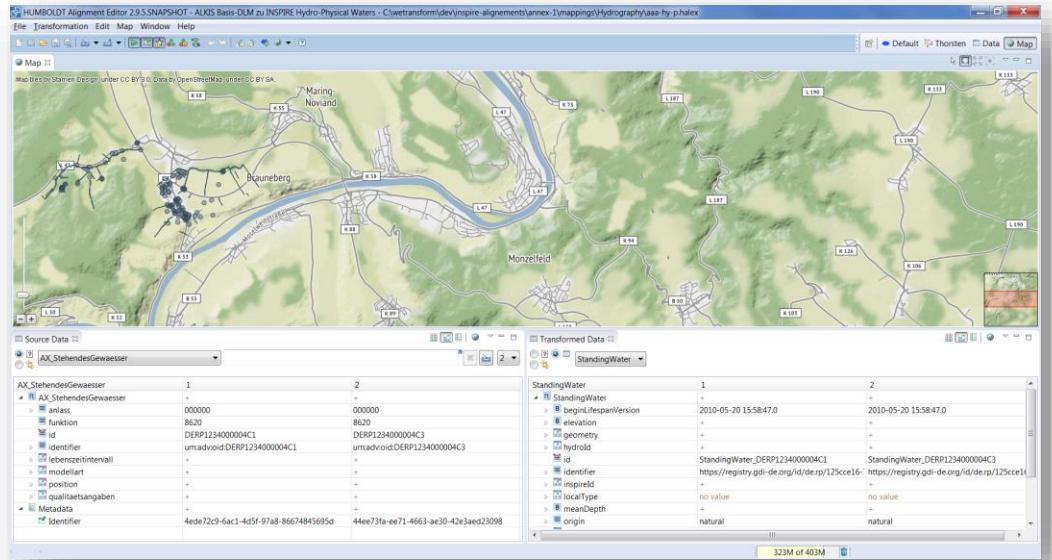
- Mismatch Data <> GIS
- Mismatch current data users <> future data users

Potential Solutions:

- Provide Libraries to create news types of visualisations from INSPIRE data
- Fully leverage Linked Data and Object Oriented Data
- Provide Adaptive Data Services and other APIs to make data accessible



Why wetransform?



- „Make Open Standards work“
- Improve Design and Implementation of Standards
- Effective Data infrastructures
- Next Generation of ETL tools
- Remove unscalable gatekeepers



Accept and
Drive Change!