Challenges and issues of data transformation from heterogeneous sources in the example of creating an INSPIRE harmonized federal waterway network

Christian Schwartze
codematix GmbH

INSPIRE Conference 2017 | Wednesday, September 6th
The “Who” question...

Informationstechnikzentrum Bund* (ITZBund, Ilmenau, Germany)

- IT service provider for mainly federal agencies
- hosting and operation of IT administrative processes, SW development and support
- initiator of this project (and other INSPIRE-related tasks)

* German Federal Centre for Information Technology

codematix GmbH (Jena, Germany)

- dynamic company specialized on customer tailored SW development, web services and training
- IT services: system design, implementation and optimization of GIS enterprise solutions; geospatial data processing, infrastructures and support
WSV

- Wasserstraßen- und Schifffahrtsverwaltung des Bundes*
- executive agency of the Federal Ministry of Transport and Digital Infrastructure
- owner and operator of the German federal waterways
- responsible for ensuring safe flowing, efficient shipping traffic and INSPIRE compliant data processing and provision

* Federal Waterways and Shipping Administration
Initially: A feasibility study

- (open source) software approaches and geodata transformation concepts for implementing the INSPIRE rules regarding Water Transport Networks (TN-W, Annex I)
- consideration of existing but heterogeneous data sources provided by various data holders
- easy extensibility (e.g. for additional data)
### Result: Types of data preprocessing

1. **XSLT**
   - eXtensible Stylesheet Language Transformation
   - Transforming XML documents into other documents or formats (rule-based)

2. **OGR**
   - OpenGIS Simple Features Reference Implementation (hist.)
   - Tools for querying, processing and manipulating vector data – such as `ogr2ogr`

3. **JAVA**
   - Object-oriented programming language
   - Using in combination with geospatial libraries – such as `Geotools`
### Existing data sources

#### (A) “VerkNet”
**ESRI Shapefile/Geodatabase**

The German transport network of inland waterways - provided by Federal Institute of Hydrology (BfG)

#### (B) Inland-ENC
**S-57 ENC chart data (.000 files)**

Inland Electronic Navigational Charts by the WSV

#### (C) “WADABA”
**Web Feature Service (WFS)**

Waterway database by the WSV
Data sources (1)

“VerkNet”

- technical complex due to professional information and agreements by the WSV (Wasserstraßen- und Schifffahrtsverwaltung des Bundes)
- diverse, position-related, detailed attributes
- example: distinction of secondary waters and main waters - geometrical delimited by connection lines

https://www.wsv.de/service/karten_geo Informationen/verknet_bwastr/index.html
Data sources (2)

**Inland-ENC**

- ECDIS format 2.3 (Electronic Chart Display and Information System), Version 2.3
- Variety of object-oriented designed vector data along inland waterways

Data sources (3)

“WADABA”

- containing facility sites or constructions along waterways that are either owned by the WSV or by any third party
INSPIRE Feature concept: Nodes and Links

**InlandWaterway**
Inland continental waterway as a collection of individual WaterwayLinks or n WaterLinkSequences

**WaterWayLink**
Two connected WaterNodes

**WaterLinkSequence**
Ordered collection of individual WaterwayLinks representing a continuous water network path without branches

**FerryCrossing**
At least one WaterwayLink across the water body

**PortArea**
Boundary of inland port

**WaterNode**
Junction or point representation of a lock, ship lift, turning basin, terminal or movable bridge

**PortNode**
Special WaterNode for point representation of a port
Data processing with OGR/XSLT (1)

**WaterNodes/-wayLinks**

- breaking down into main relevant components: individual junction and waterway GML files
- as input for Java processing

Example:
- junctions_3204.gml
- verknet_3204.gml
Data processing with OGR/XSLT (2)

WaterNodes

- retrieving relevant facilities from Web Feature Service (WFS)
- requests for locks, ship lifts, turning basins, (movable) bridges
- as input for Java processing

```
ogr2ogr.exe -f "ESRI Shapefile"
-s_srs "EPSG:4326"
-t_srs "EPSG:900913"
-append -skipfailures
-nln result
-select %OBJEKTE_WADABA_SELECT%
"..\data\tmp\wadaba_objekte" WFS:%WADABA_WFS_URL% %WADABA_LAYER%
--config GML_INVERT_AXIS_ORDER_IF_LAT_LONG YES
--config GML_CONSIDER_EPSG_AS_URN YES
```

Resulting attribute table of INSPIRE relevant WADABA objects
Data processing with OGR/XSLT (3)

**PortArea/FerryCrossing**

- querying IENC data files for relevant S-57 objects, e.g. "hrbbsn" and "feryrt"
- as input for Java processing

---

**Resulting attribute table of port objects**

```
ogr2ogr.exe -f "ESRI Shapefile"
   -s_srs EPSG:4326
   -t_srs EPSG:900913
   -update -append
   -nln %IENC_LAYER_NAME%
   %IENC_PATH%/MERGED/%IENC_LAYER_NAME%.shp "%g"
   %IENC_LAYER_NAME% --config S57_PROFILE iw
```
Geodata processing with Java (1)

**WaterNodes/-wayLinks**

- `verknet_3204.gml`: connection lines (red) not desirable/not needed in INSPIRE context
- create **continuous lines** from one junction to the other
- merging connection lines with waterway line
- use junctions from e.g. `junctions_3204.gml`
Geodata processing with Java (2)

WaterNodes/-wayLinks

- geometries in data sources are “not harmonized”
- merging waterway lines and water nodes needed
- waterway-by-waterway (e.g. verknet_3204.gml): spatial index searching for candidates (e.g. locks, ship lifts, ports…) from previous WADABA Shapefile
- ordering the hits and snapping them along the waterway
- final inserting
Geodata processing with Java (3)

WaterNodes/-wayLinks

- ordered, snapped and final water nodes split current waterway into parts
- results in new INSPIRE waterwaylinks with start/end node
- sequential ID numbering

<table>
<thead>
<tr>
<th>INSPIRE Feature Type</th>
<th>INSPIRE IDs (without namespace)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WLS WaterLink-Sequence</td>
<td>e.g. WLS-6825.1.006</td>
</tr>
<tr>
<td>WWL Waterway-Link</td>
<td>e.g. (A) WWL-6825.1.005.001 (B) WWL-6825.1.006.001 (C) WWL-6825.1.006.002 (D) WWL-6825.1.007.001</td>
</tr>
<tr>
<td>WND Waterway-Node</td>
<td>WND-LOC-6825.1-0003193.32 WND-JUN-6825.1-0004834.54</td>
</tr>
</tbody>
</table>

Minimalist example at federal waterway 6825 “Flussstrecke Rathenower Havel”
Geodata processing with Java (4)

WaterNodes/wayLinks

- ID schema used and valid in all GML objects delivered by *deegree* software
  
  [http://www.deegree.org](http://www.deegree.org)
Geoata processing with JAVA (5)

**PortNodes**

- INSPIRE: „A point [...] to represent a sea or inland port in a simplified way, approximately located at the bank of the waterbody [...]”
- uses Ports Shapefile from OGR preprocessing
  - identify centre point of harbour polygon and connect it with waterway (shortest distance) → **new point** is new junction!
  - find shortest way between new point and polygon → **new point** is port representation (PortNode)!
Resulting INSPIRE WMS for TN-W
Thank you!