COMBINING ATTRIBUTE DATA AND SPATIAL DATASETS WITH THE UPDATED TABLE JOINING SERVICE STANDARD

Pekka Latvala
INSPIRE Conference 2018
18-21.9.2018
Antwerp, Belgium
TABLE JOINING SERVICE

- Table Joining Service (TJS) 1.0.0 standard has been defined by the OGC in 2010
- Main features
  - Describing and exchanging tabular attribute data that relates to geographic objects
  - Describing metadata on spatial datasets (i.e. post area codes, municipalities, etc…)
  - Joining the attribute data with different spatial datasets
- Uses a XML-based GDAS format for describing attribute data
- The data joining is executed through common geographic identifiers that are shared between the spatial and attribute datasets
TJS UPDATE – TJS 2.0

- TJS update work has been started in the OGC in the spring 2018,
- The work is currently ongoing
- Update plan
  - Define TJS 2.0 as a RESTful service
  - Interaction with the service through HTTP queries
    - GET, POST, PUT, DELETE etc..
  - JSON as main output format instead of GDAS
  - OpenAPI specification for interface description
    - JSON or YAML format
    - Replaces GetCapabilities
    - Can be visualized for example with Swagger UI
IDENTIFIED CHANGE REQUESTS

1. More flexibility to input and output formats
   - More supported formats/services for both tabular and geospatial data i.e. CSV, Excel, SDMX, JSON-stat, WFS, GeoJSON

2. Enhancements to data joining
   - Metadata for the Join operation (Unmatched ID's, number of mismatches)
   - Case insensitivity in data joining.

3. Data handling in TJS
   - Normalization (area, population, other)
   - Different calculations: Calculate the mean or the sum, text concatenation etc…

4. Persistent ID's and mapping enhancements
   - Use of persistent ID's for all references to geospatial features & tabular data.
   - ID mappings - map IDs between services
TJS UPDATE – STRUCTURE

- Plan to define TJS 2.0 specification with a modular structure

Mandatory modules

CORE MODULE

Attribute data and metadata and data viewing and retrieval
Spatial dataset metadata viewing
Data joining

Possible extension modules

ADMINISTRATOR MODULE

Attribute dataset & spatial dataset adding / updating / deleting

ID MAPPING MODULE

ID mapping

DATA HANDLING MODULE

Data normalization
Data aggregation

OTHER MODULES
CURRENT STATUS

• TJS service demo implementation is being developed in parallel to the specification work
• Current operations:

```plaintext
attributedatasets
GET /attributedatasets Get metadata of all attribute datasets
GET /attributedatasets/{datasetid} Get metadata of specified attribute dataset
GET /attributedatasets/{datasetid}/data Get data of specified attribute dataset
```

Data filtering through query parameters

```plaintext
joinabilities
GET /joinabilities Get metadata on server's output abilities
```

```plaintext
joindata
GET /joindata/{spatialdatasetid} Join attribute dataset to a spatial dataset
POST /joindata/{spatialdatasetid}/csv Join attributes from csv file
```

+ query parameter attributedataseturl
DEMO SERVICE

- Demo service that joins CSV data and spatial datasets
- User selects a spatial dataset available on server
- User inputs an CSV file
  - Key values must match in both datasets
- User inputs information about the CSV file
  - Title line number
  - data start line number
  - key column number
  - data column numbers
  - column separator character
**DEMO SERVICE**

- Service joins the datasets
- The joined dataset is published as a layer to GeoServer
- The demo service offers multiple outputs
  - GeoJSON
  - Shapefile
  - CSV
  - KML
- Information about the successfullness of the join operation
  - Matches, Mismatches
  - Additional keys in input CSV file
INTERESTED?

- TJS 2.0 definition work is currently ongoing in the OGC
- If you are interested, come to OGC meetings and join the SWG work
- Give feedback and comments, ideas
- Participate in GitHub https://github.com/opengeospatial/tjs
THANK YOU