On-demand mapping and integration of thematic data
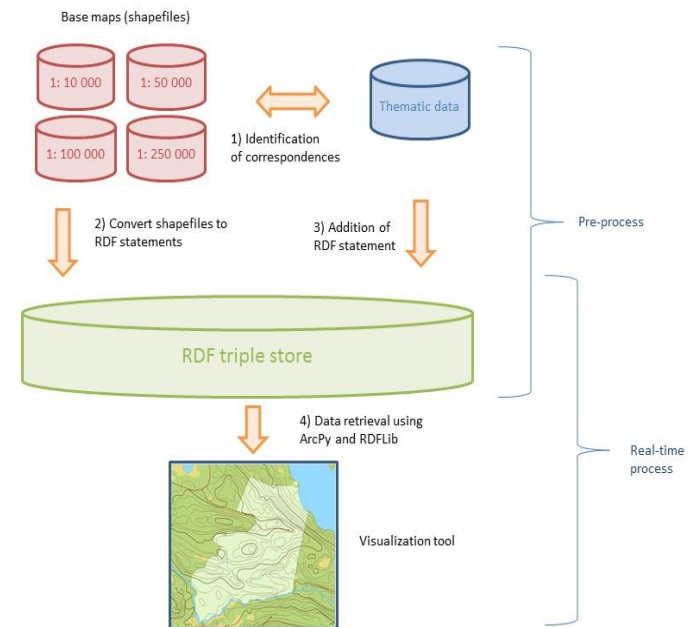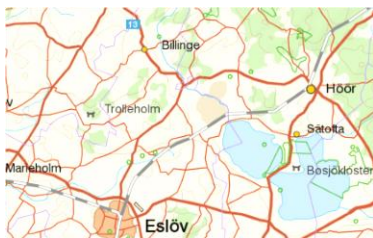
Weiming Huang

Ali Mansourian

Lars Harrie

Department of Physical Geography and Ecosystem Science

Lund University

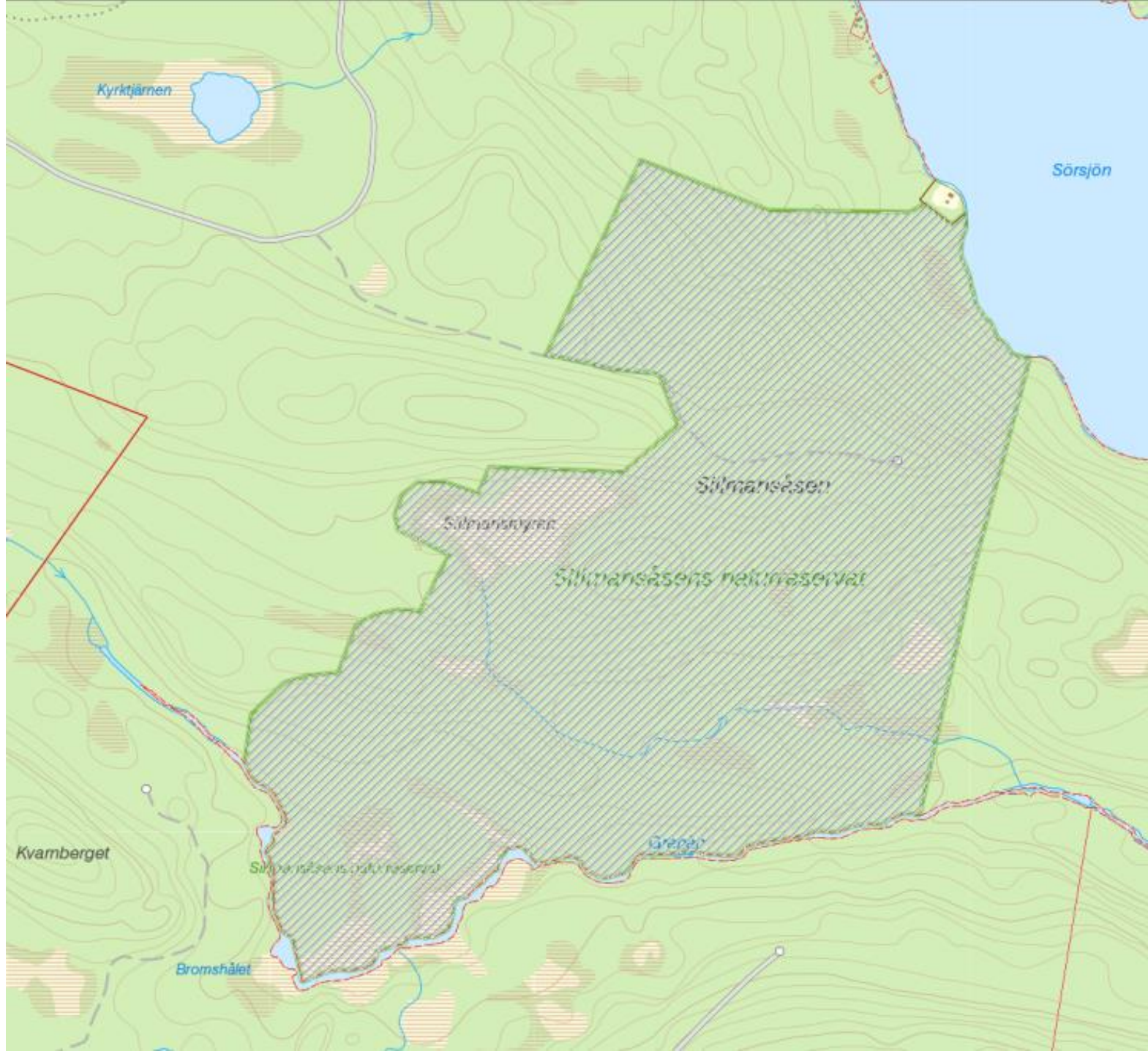# Utilising a multiscale base maps to provide thematic data in several LoD

Scale

Base map

Thematic data

LUND
UNIVERSITY

Kyrktjärnen

Sörsjön

Silmansåsen

Silmanstjärn

Silmansåsens naturreservat

Kvamberget

Silmansåsens naturreservat

Grepen

Bromshålet

Linked Geodata and RDF

Linked geodata

- Geodata where you stored explicit links to model relationships between features. Data often stored in RDF.

RDF – Resource Description Framework

- All data are stored as triples: a subject, a predicate and an object, e.g.:

lake_idi_2 *hasGeometry*   *Geometry (e.g. stored as OGC WKT)*

- Structure of triples are defined in ontologies e.g.:
Geometry ontology and *Spatial Relations Ontology*

Base maps distributed as linked data

OS Ordnance Survey

Ordnance Survey Linked Data Platform

| Quick Search: | Enter a place name or postcode to view as | 🔍 |

Ordnance Survey is Great Britain's national mapping agency, providing the most accurate and up-to-date geographic data, relied on by government, business and individuals.

OS OpenData is the opening up of Ordnance Survey data as part of the drive to increase innovation and support the "Making Public Data Public" initiative. As part of this initiative Ordnance Survey has published a number of its products as Linked Data. Linked Data is a growing part of the Web where data is published on the Web and then linked to other published data in much the same way that web pages are interlinked using hypertext.

The term Linked Data is used to describe a method of exposing, sharing, and connecting data via URIs on the Web. To find more Linked Data published as part of this initiative please go to data.gov.uk.

If you are not familiar with Linked Data, OS OpenData products are also available in alternative formats from the OS OpenData website. Ordnance Survey can provide support for the Ordnance Survey OpenData products, but cannot give advice or support on using RDF, SPARQL or SPARQL Endpoints.

Ordnance Survey has published three OS Open Data products as Linked Data: the 1:50 000 Scale Gazetteer, Code-Point Open and the administrative geography for Great Britain taken from Boundary Line. A combined OS Linked Data dataset combines these products into one database to support more flexible data access.

Each of the datasets is accessible as Linked Data and via a range of APIs.

Current Datasets
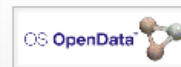
OS OpenData
Ordnance Survey Linked Data
58,233,148 triples

Code-Point Open Linked Data
33,892,896 triples

Open Names Linked Data
21,276,861 triples

50K Gazetteer Linked Data
2,362,412 triples

Boundary Line Linked Data
699,499 triples

Base maps distributed as linked data

LinkedGeoData.org
Adding a spatial dimension to the Web of Data.

Project created by:
AKSW

About / News
Downloads
Online Access
RDF Mapping
**Use Cases**
LGD Browser
Publications
Community
Blog
Contact / Imprint

2016 Jan 13: New Open Street Map RDF datasets available online and for download!
Quick Links: Downloads – SPARQL – Virtual-SPARQL by Sparqlify – HTML interface – Example Queries

LinkedGeoData is an effort to add a spatial dimension to the Web of Data / Semantic Web. LinkedGeoData uses the information collected by the OpenStreetMap project and makes it available as an RDF knowledge base according to the Linked Data principles. It interlinks this data with other knowledge bases in the Linking Open Data initiative.
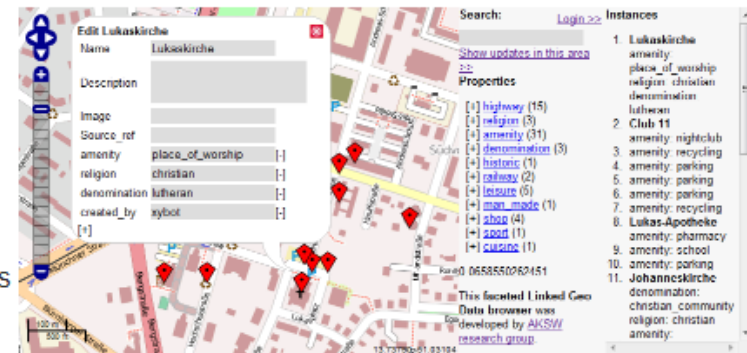
News

LinkedGeoData: New RDF versions of OpenStreetMap datasets available
The AKSW research group is happy to announce that a new LinkedGeoData maintenance release with more than 1.2 billion triples based on the OpenStreetMap planet file from 2015-11-02 is now online. Enjoy! Quick Links Project Website Downloads SPARQL Endpoint Virtual … Continue reading →
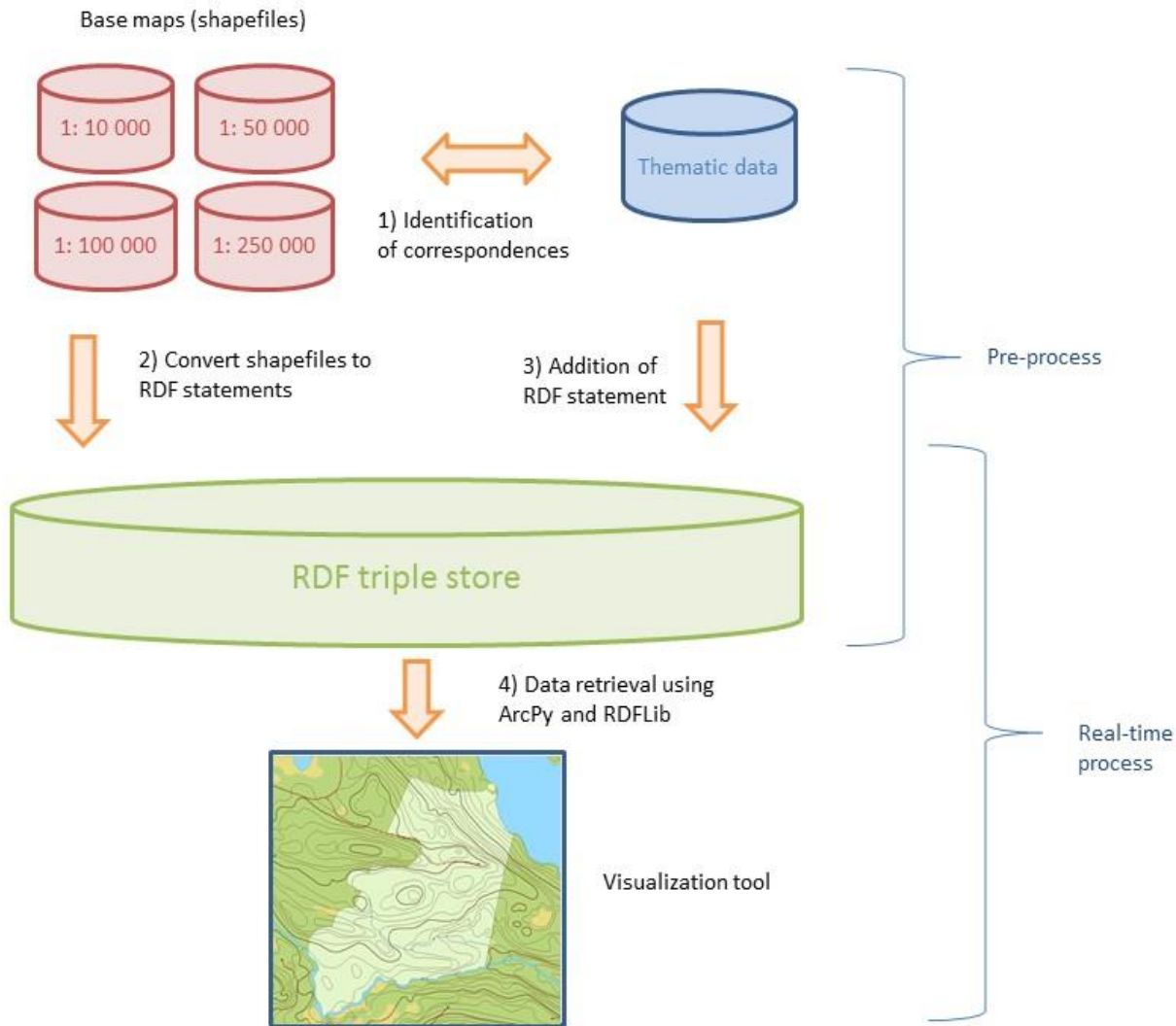
AKSW at #ISWC2014. Come and join, talk and discuss with us!
Hello AKSW Follower! We are very pleased to announce that nine of our papers were accepted for presentation at ISWC 2014. In the main track of the conference we will present the following papers: AGDISTIS – Graph-Based Disambiguation of Named … Continue reading →
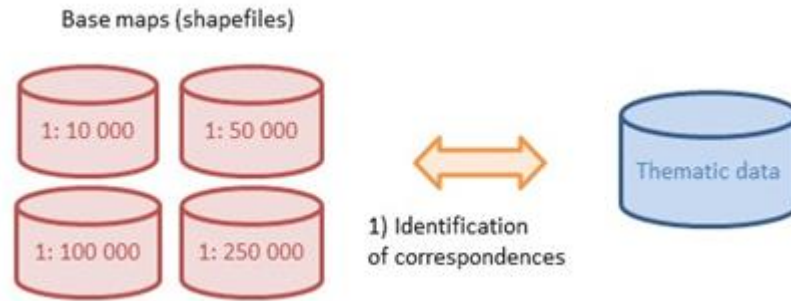
AKSW at TU Dresden PLT
On June 8, I (Jens) visited the process control engineering research group (PLT) of Leon Urbas at the Dresden University of Technology. We first met on the Leipziger Semantic Web Day where Leon Urbas presented interactive Linked Data applications and … Continue reading →

Proposed methodology

Base maps (shapefiles)

1: 10 000 1: 50 000

1: 100 000 1: 250 000

Thematic data

1) Identification of correspondences

2) Convert shapefiles to RDF statements

3) Addition of RDF statement

RDF triple store

4) Data retrieval using ArcPy and RDFLib

Visualization tool

Pre-process

Real-time process

Step 1: Identification of correspondences

Base maps (shapefiles)

1: 10 000   1: 50 000

1: 100 000   1: 250 000

1) Identification of correspondences
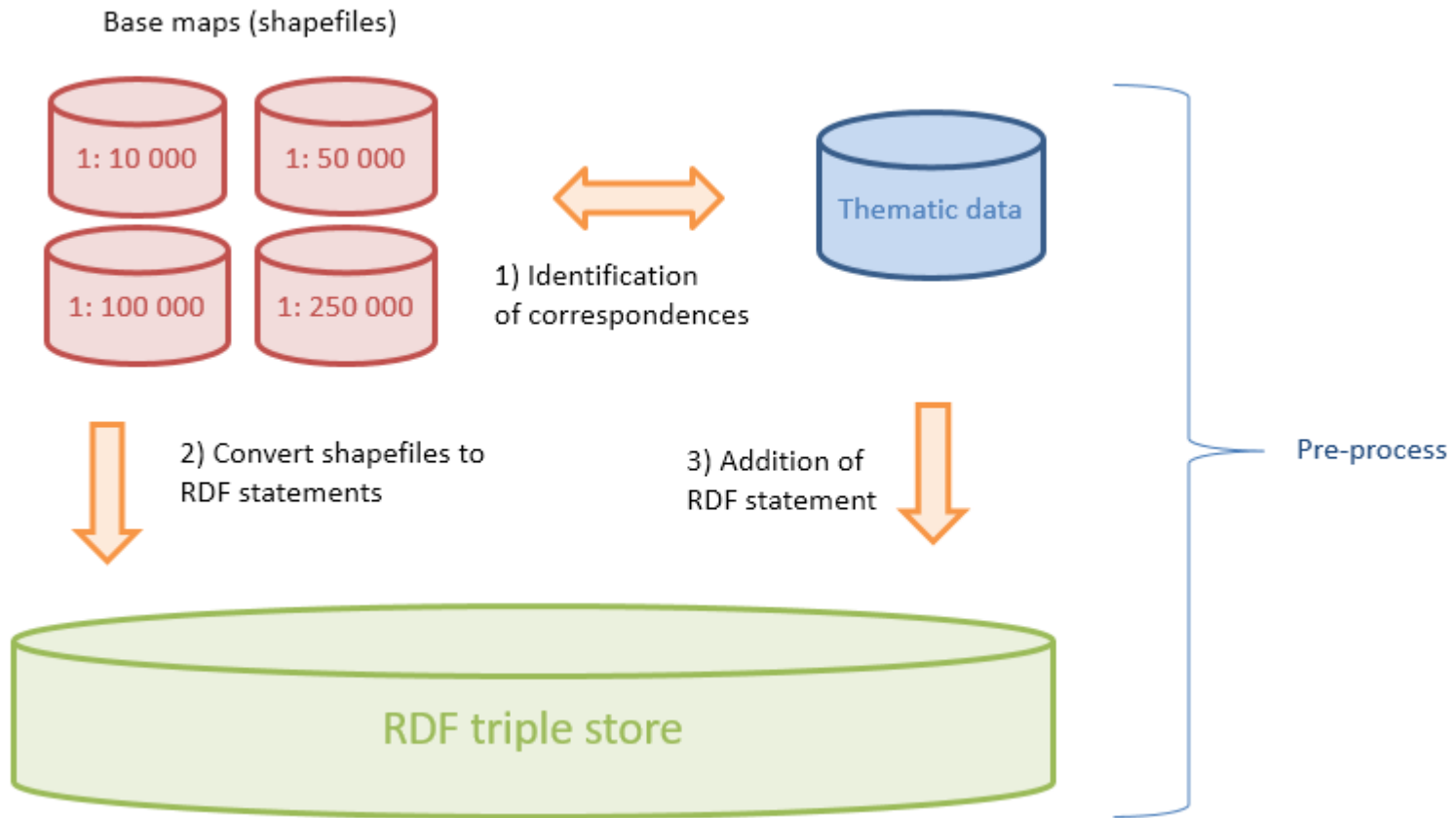
Thematic data

Data matching: conducted once.

Resulting in that a thematic feature could be linked to parts of geometry of several features on the base maps.
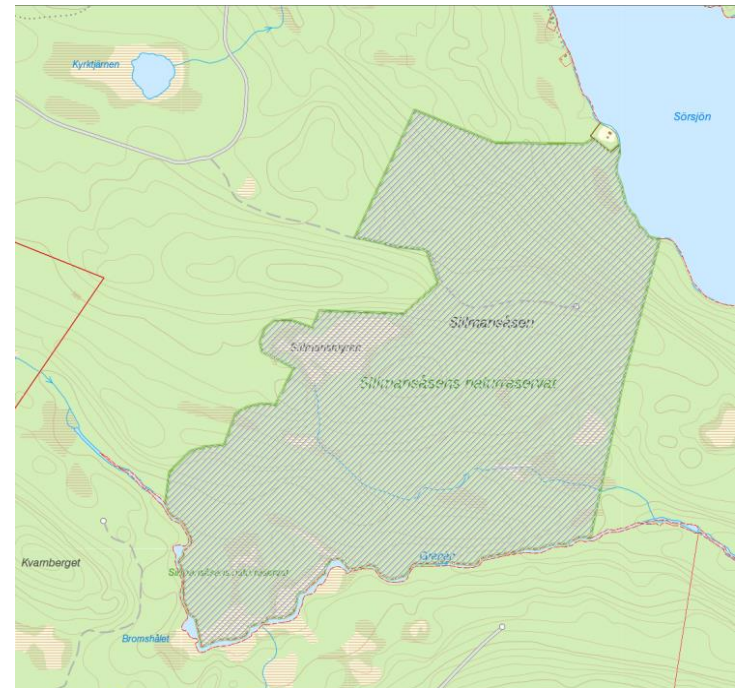
Step 2 and 3: Conversion to RDF

As to the conversion of original geographic data
For this we plan to use the RDFLib (https://github.com/RDFLib). The RDF statements are then added to a RDF triple store (most likely Sesame, see http://rdf4j.org/).

RDF triples
(created in step 2 and 3)

……

<!--Geometric information of features on base map under several scales-->

lane_idi_1	hasGeometry_50k	<coordinates>
lane_idi_1	hasGeometry_100k	<coordinates>
lane_idi_1	hasGeometry_200k	<coordinates>

……

lake_idi_2	hasGeometry_50k	<coordinates>
lake_idi_2	hasGeometry_100k	<coordinates>
lake_idi_2	hasGeometry_500k	<coordinates>

……

<!—Representing a part of the thematic feature using the combination of parts from base map features -->

| thematicFeature_idi_3 | hasComponent | subThematicFeature_idi_3_1 |
subThematicFeature_idi_3_1	isPartOf	lane_idi_1
subThematicFeature_idi_3_1	startsAt	<coordinate_of_one_point>
subThematicFeature_idi_3_1	endsAt	<coordinate_of_one_point>
thematicFeature_idi_3	hasComponent	subThematicFeature_idi_3_2
subThematicFeature_idi_3_2	isPartOf	thematicFeature_idi_3
subThematicFeature_idi_3_2	startsAt	<coordinate_of_one_point>
subThematicFeature_idi_3_2	endsAt	<coordinate_of_one_point>

……

Step 4: Data retrieval and visualisation
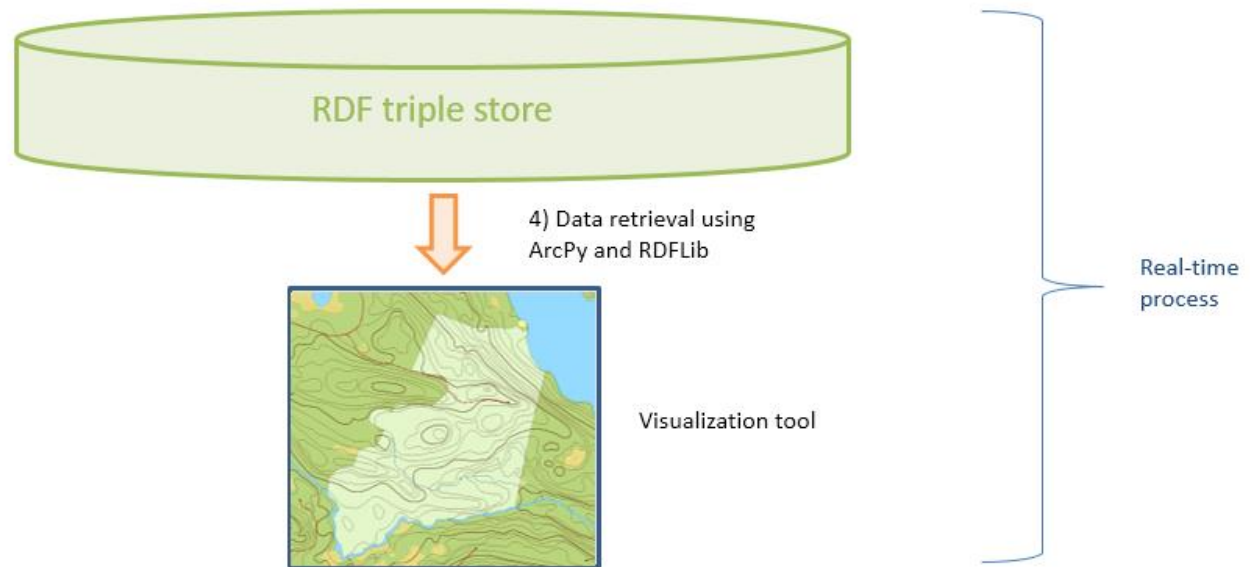
- **The visual geometry of the thematic features is constructed by a combination of the parts of the counterparts on the base maps, and parts remained from the thematic data.**

- **Data retrieval: querying the triple store using SPARQL**

- **The manipulation and visualization: ESRI Python programming interface ArcPy.**

Concluding remarks

It is:

- **A way to represent geographic thematic features**

- **A way to put the thematic data into a multi-scale world**

Advantages:

- **Lifting the visual performance in cartographic mashups**

- **As a result, the data are lifted to the Semantic Web infrastructure which can benefit further integration**

Disadvantages:

- **Perhaps quite time-consuming during both the pre-process and real-time process**

- **Defining proper ontologies would be tricky.**

LUND
UNIVERSITY

Thank you for your attention!