



IGALOD Project

(Integration of Geographies and Areal
Classifications as Linked Open Data)

Tuuli Pihlajamaa STAT-FI, Kai Koistinen NLS

Nordic Forum for Geography and Statistics 14th-15th May 2019

Abstract

1. IGALOD Project

2. DEMO

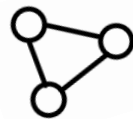
3. Next steps



1. IGALOD Project

What is IGALOD?

- 2-year cooperational project between Statistics Finland (NSI) and National Land Survey of Finland (NMA) (2018-2019)
- Funded partly by Eurostat Grant Program
- Linking areal classifications and corresponding geographies to:
 - Publish them as Linked Open Data (LOD)
 - Utilize them in statistical production process
 - Use them to create interactive map applications

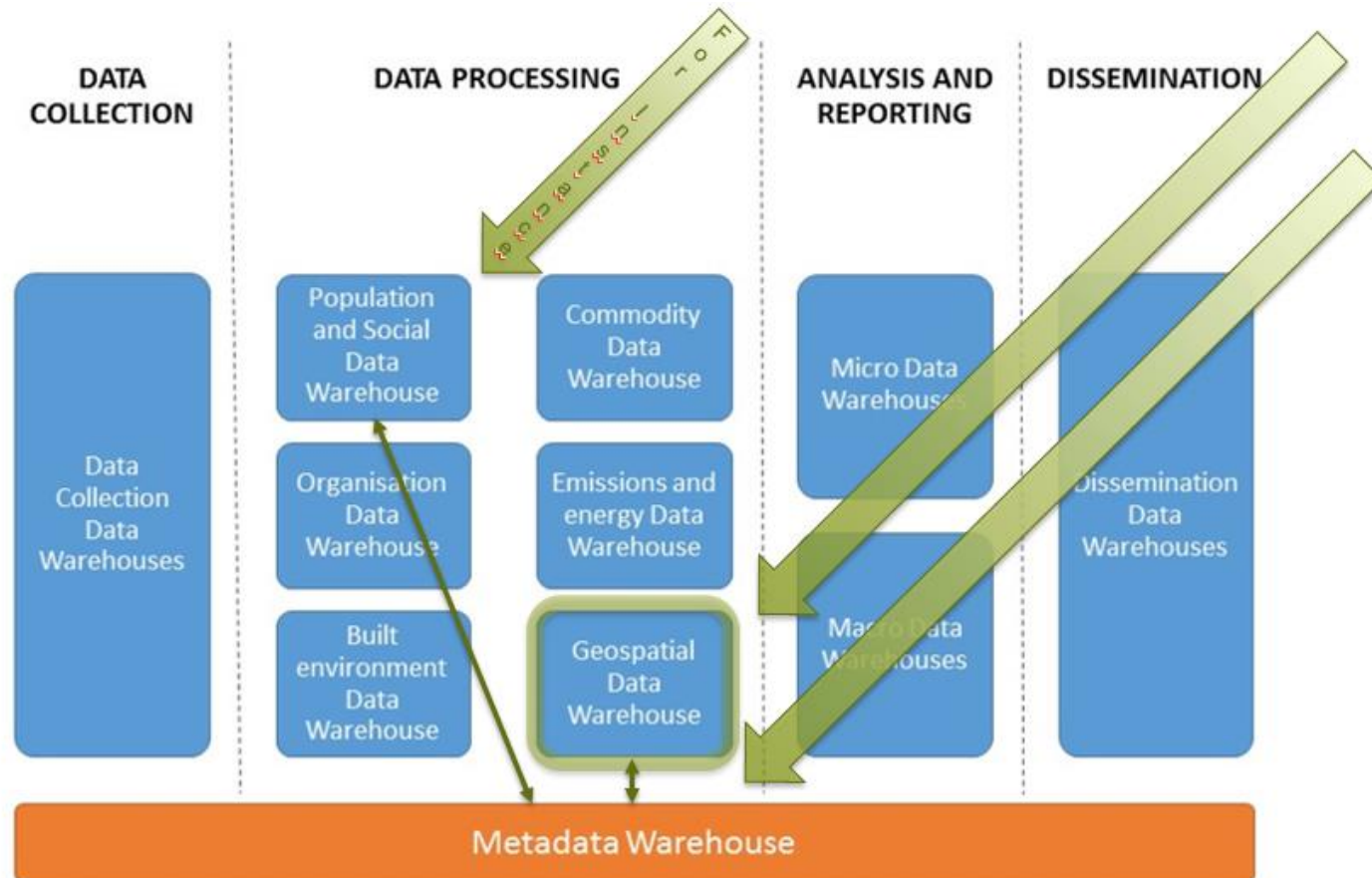


GSBPM

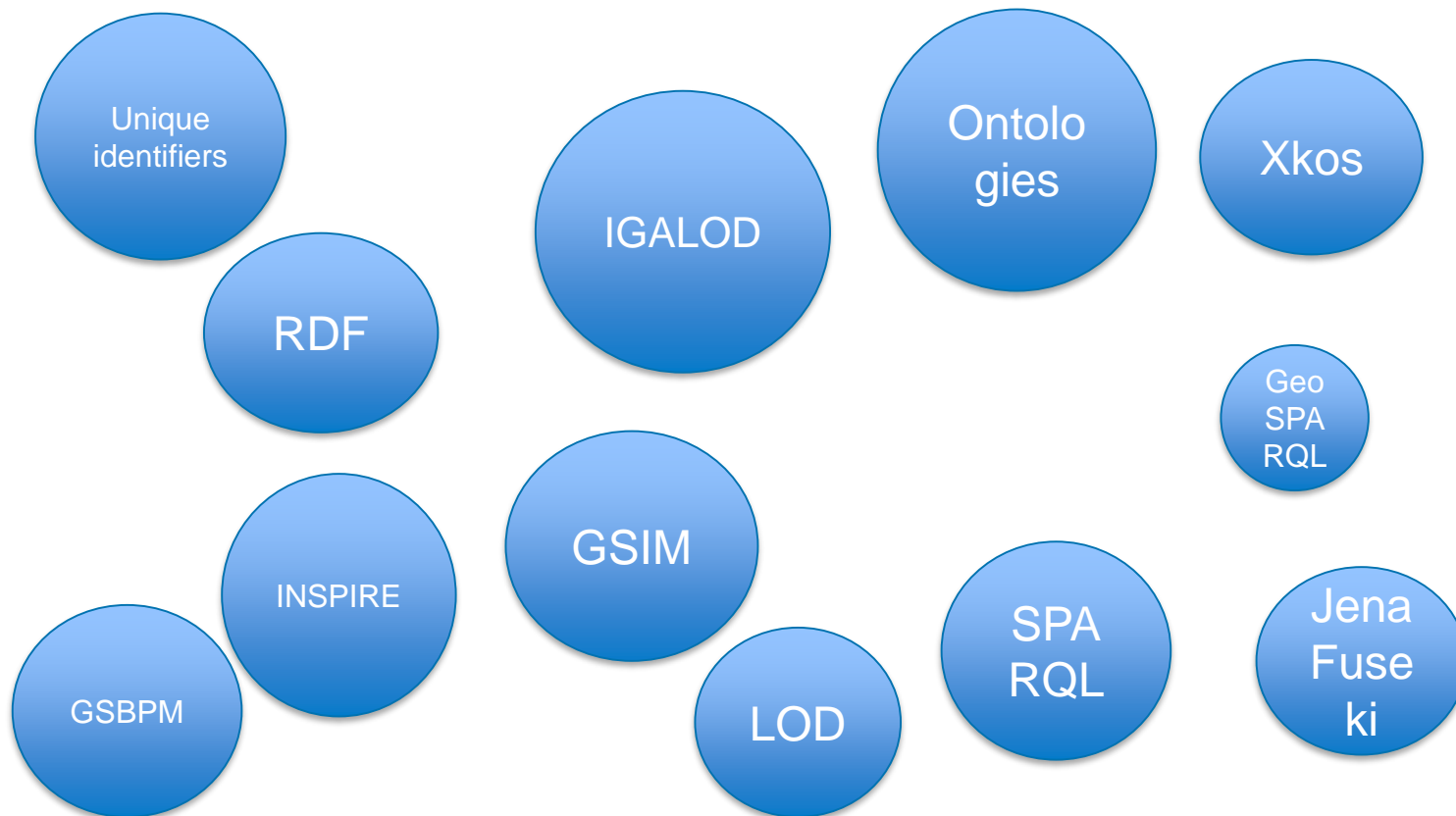
Expected outcomes

- Identifier system, identifiers and required ontologies
- RDF-based solution, incl. architecture description
- Piloting the LOD dissemination & interactive map applications
- Cross-organizational production process for geospatial information, utilizing Linked Data

Big picture: Linking GSIM statistical area classifications and corresponding geographies



Terms and standards



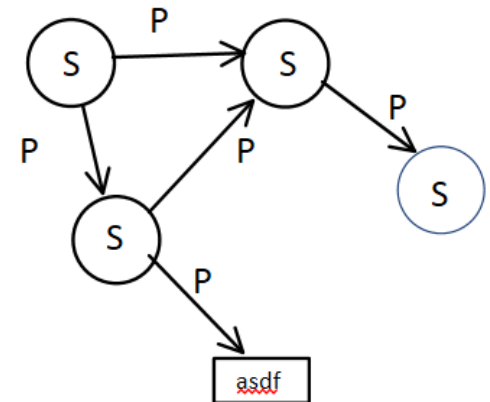
What is RDF and why do we use it?

- Abstract data model – can work on top of different syntaxes and on different databases
- All information will be expressed on sentences (triples) with three parts: *subject*, *predicate*, *object*.
 - Each part of sentence (typically) have their own URI

Graph

- All relations are possible, also outside
- Meanings are expressed as sentences
- Linked data, networks

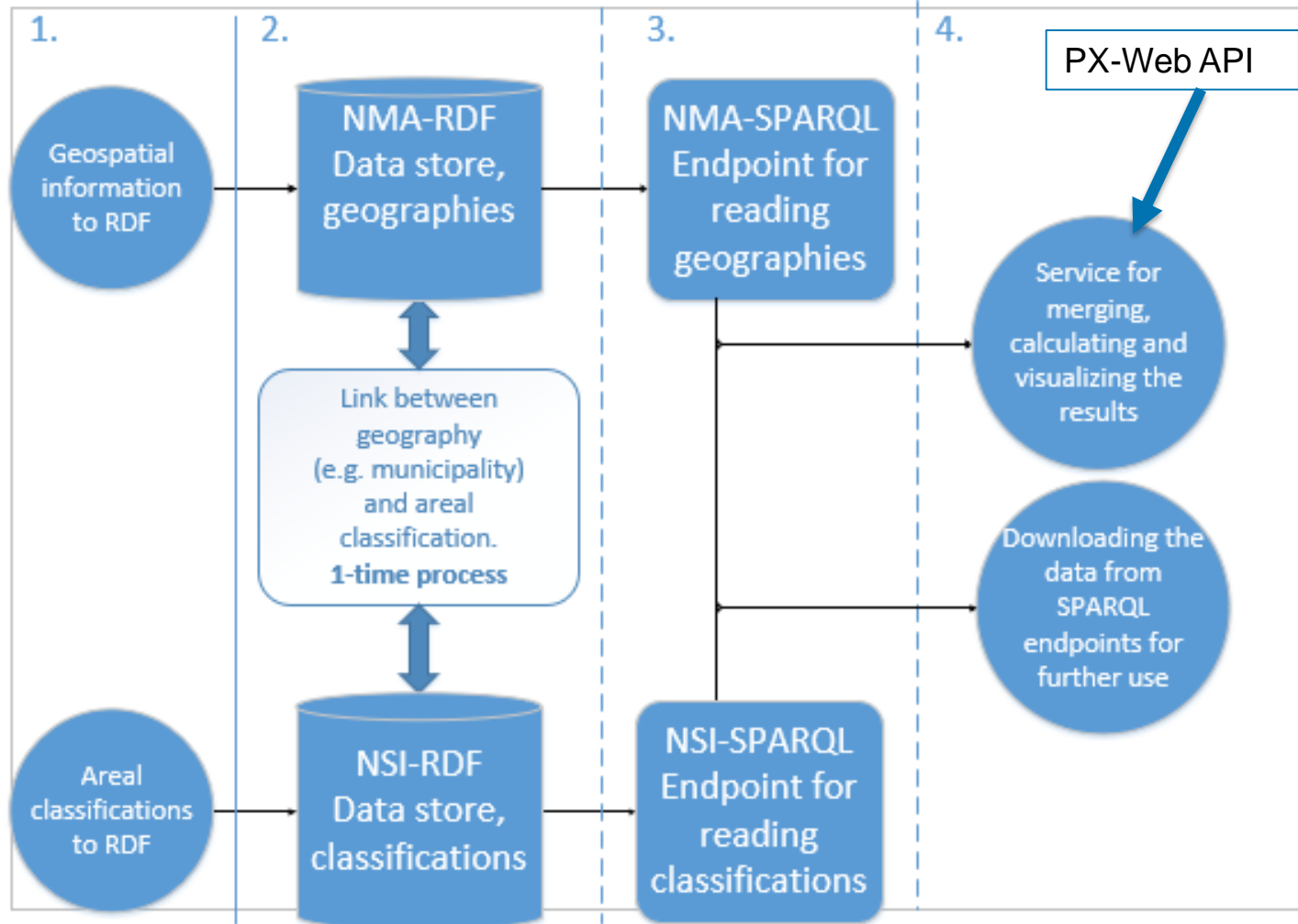
→ Good for metadata in machine-readable format!



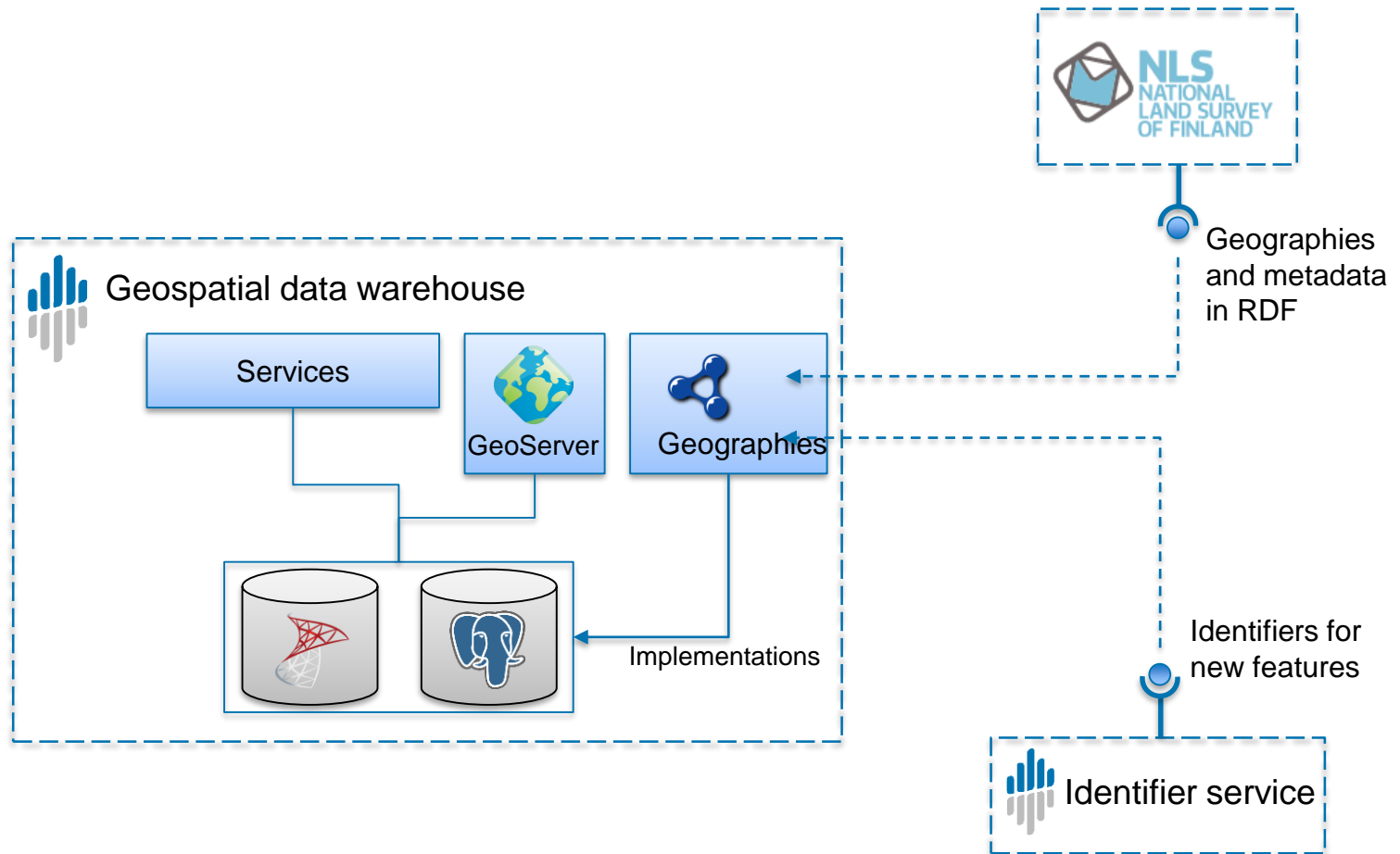
Standards

- **RDF of classifications →**
 - From RDF vocabularies XKOS was chosen
 - XKOS is SKOS extension for representing statistical classifications
- **RDF of geographies →**
 - INSPIRE AU ontology
 - GeoSPARQL ontology
 - ...and some own non-standardized concepts

Mapping architecture



"STINA architecture" (to be tested)





2. DEMO

Demo service and data

- Link to SPARQL demo service is here:
<http://193.167.189.160/igalod/fuseki/>
- Both geographies, classifications and some sample statistics was downloaded to the demo service
- In the final service model NLS will provide their data from the same server, but STAT-FI will soon start their own service
→ Data can be queried from two services using federated query

Demo queries

- Querying combination of geographic data, an areal class (municipality) and statistical data. (<http://yasgui.org/short/E7rL2s5sg>)
- Querying combination of an areal class (Region), included municipality classes according to classification key, corresponding geographic data and corresponding statistical data. (<http://yasgui.org/short/sB98izjwM>)



3. Next steps

Visualisation service

- Ordered from subcontractor
 - First phase with basic functionalities in June
 - Second phase with more functionalities in fall
- Source code will be released as Open Source code

First phase: planned functionalities

- Municipalities or selected areas (e.g. Regions) on map and statistics of the areas on table
 - On what other area classifications the municipalities belong to
 - What municipalities belong to the selected area/class
- Calculate sum aggregates and density values for selected classifications
- Select municipalities to create own areas and name the areas
 - See summed values of indicators of selected municipalities

Second phase: planned other functionalities

- Give the edge value of indicator and based on that there will be created new areas from municipalities next to each other
 - Edit the calculation algorithm in window view
 - See the summed statistics and the density values on the new areas
- Option for saving and downloading the results

Data used

- Geographies
 - NLS provides from their SPARQL endpoint
 - Includes also the official surface areas in m² as metadata
- Classifications, classification keys
 - STAT-FI provides from their SPARQL endpoint
- Statistics
 - STAT-FI provides from Px-WEB API

Used geographies and classifications

- Geographies from 2018 and 2019
 - 1:1 000 000 and 1:4 500 000, both with and without sea areas
 - Also 1:10 000
- Classifications from 2018 and 2019
 - Regions (NUTS2016)
 - Municipalities (LAU 2019)
 - Major regions
 - Sub-regional units
 - Constituencies 2019
 - Statistical grouping of municipalities
 - Travel-to-work areas
 - Degree of Urbanisation (DEGUBRA)

Used statistics

- Population
 - Number of families
 - Number of household-dwelling units
 - Employed labour force resident in the area
 - Number of workplaces in the area
-
- https://pxnet2.stat.fi/PXWeb/pxweb/fi/Kuntien_avainluvut/Kuntien_avainluvut_2018/kuntien_avainluvut_2018_viimeisin.px/?rxid=da8ad535-2ec8-4d57-af2a-f9c2ea0f4a10



Thank you!

Statistics Finland 