

# Low hanging fruits and forward with Spatial data on the web

Spatial Data on the Web –  
Web Linked Data initiatives&issues from Finland

Esa Tiainen

Spatial Data on the Web Interest Group – SDWIG

Toulouse 19.11.2019



# National Land survey of Finland

- National mapping, dwelling and cadastral authority (NMCA )
- Also national Land registration authority
- Responsible of NSDI implementation
- **Strategic goal: Interoperability**
- Working to introduce a URI-based management of SDI
  - A national recommendation for public administration on unique HTTP URI identifiers for SDI, approved 2015
  - Developing URI/RDF-based Framework for Linked SDI
  - Current implementation: URI-based maintenance, data linking and delivery in the new geospatial platform(s);
  - semantic feature search by attributes;
  - next: machine readability and APIs
-

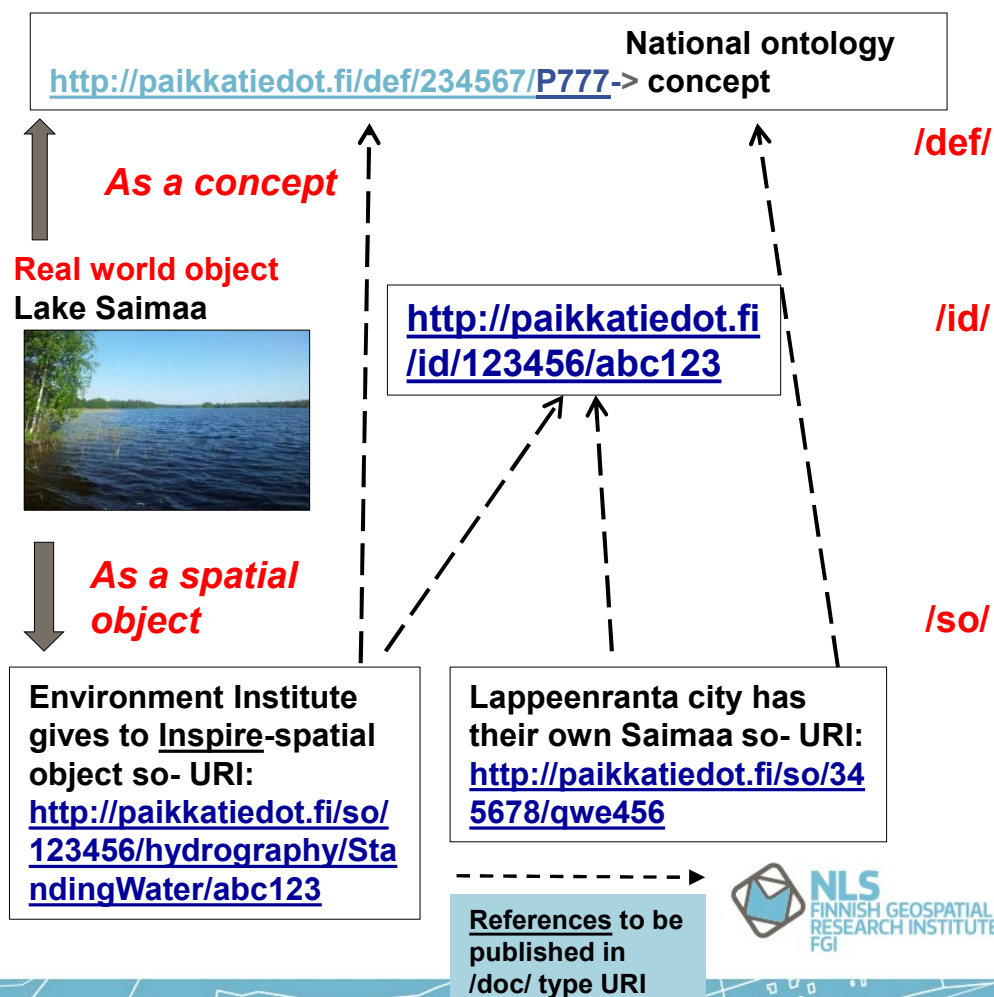
# Content

- National URI-recommendation on spatial data & data linking principles
- New national geospatial platform (NSDI)
- Towards nationwide linked data and SDI
- Low hanging fruits for Linked data (SDI)
- Next: Machine readability, 'Geographic space'
- Georef – data linking by placenames
- UN SDG related – Agenda30

# URI recommendation (2015) & linking principles of SDI

4

- <http://{domain}/{type}/{datasetId}/{localId}/{versionId}>
- type /so/= spatial object, but publishing http URIs also
- For concepts to link spatial objects with concepts - */def/*
- For real world entities a *placeholder-URI* to enable combining data and searches of spatial objects modeling the same real world entity - */id/*
- Suggested referencing from spatial objects to real world entities and concepts, which are embedded in national ontology service



# Towards nationwide linked data SDI

- The new geospatial platform is directing to nationwide LD of spatial data
  - By combining spatial data from different providers
  - By establishing a new linked data infrastructure
  - By linking data via real world objects (spatial things)
  - By enabling to integrate spatial and non-spatial data via real world objects or concepts
- Current stage: First on buildings, administrative units, hydrography, road network, land cover (pilot applications) – INSPIRE aligned
- Next release: semantic feature search by attributes;
- next: machine readability and APIs

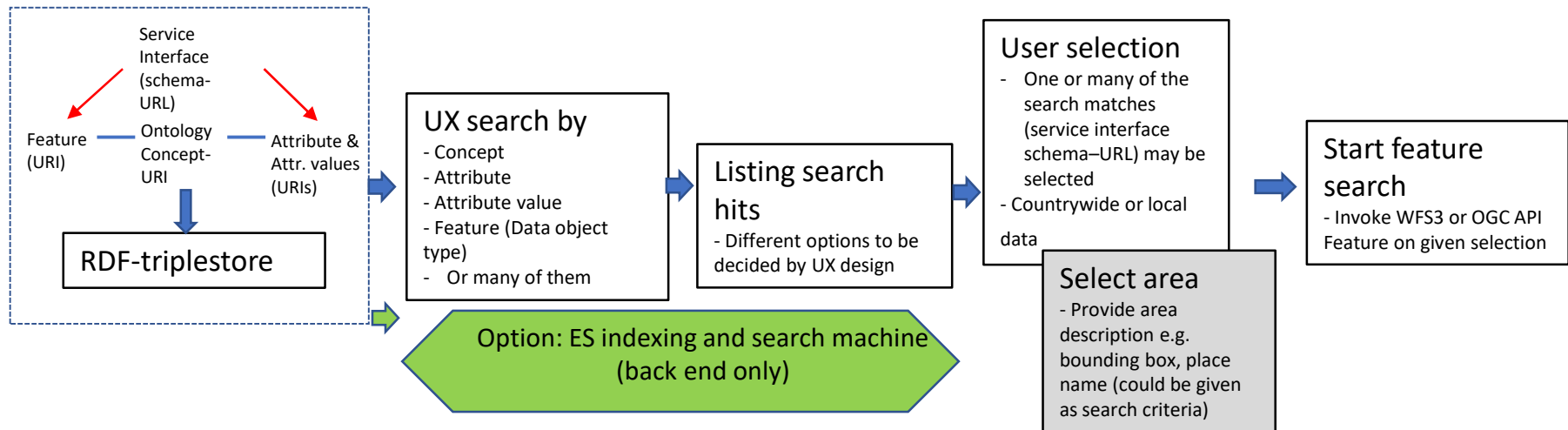
# low hanging fruits - Nordic NMAs

## Linked data group (NOSIN LD group)

- Identified the two low hanging fruits to deploy LD/Web APIs
    - Semantic feature search by attribute filtering
    - HTML-cards of features/spatial things available
  - To enable wider user take-up
  - To provide access to data through web
- 
- The idea here is to **define successive stages in web deployment** in a way that at every stage more benefits are possible and can be verified to justify the costs

# Structuring semantic feature search with attribute filtering in the new National Geospatial Platform of Finland

## Overview



Features, attributes and attribute values are bridged to the concepts that they present.

Each entity in service interface schemas (features, attributes and attribute values) are annotated to their URIs.

UX design specifies the grouping of hits as regards features, attributes or their values (with one or many search terms)

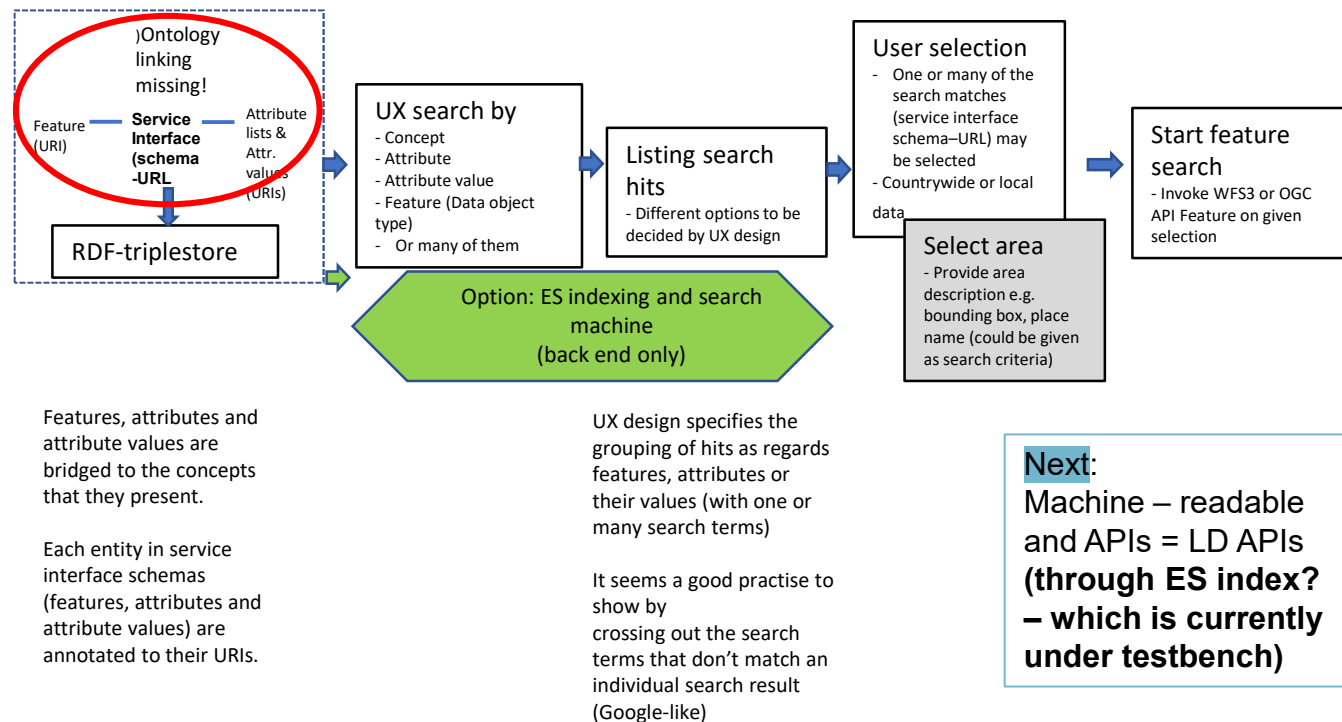
It seems a good practise to show by crossing out the search terms that don't match an individual search result (Google-like)

**Next:**  
Machine – readable  
and APIs = LD APIs

# Structuring semantic feature search with attribute filtering in the new National Geospatial Platform of Finland

First Implementation – Going live and open 4/2020:

Because of missing an ontology with adequate granularity as for natural or common language search, we had to turn to a faceted search – an alphabetical drop-down bar with some fuzzy taxonomy!



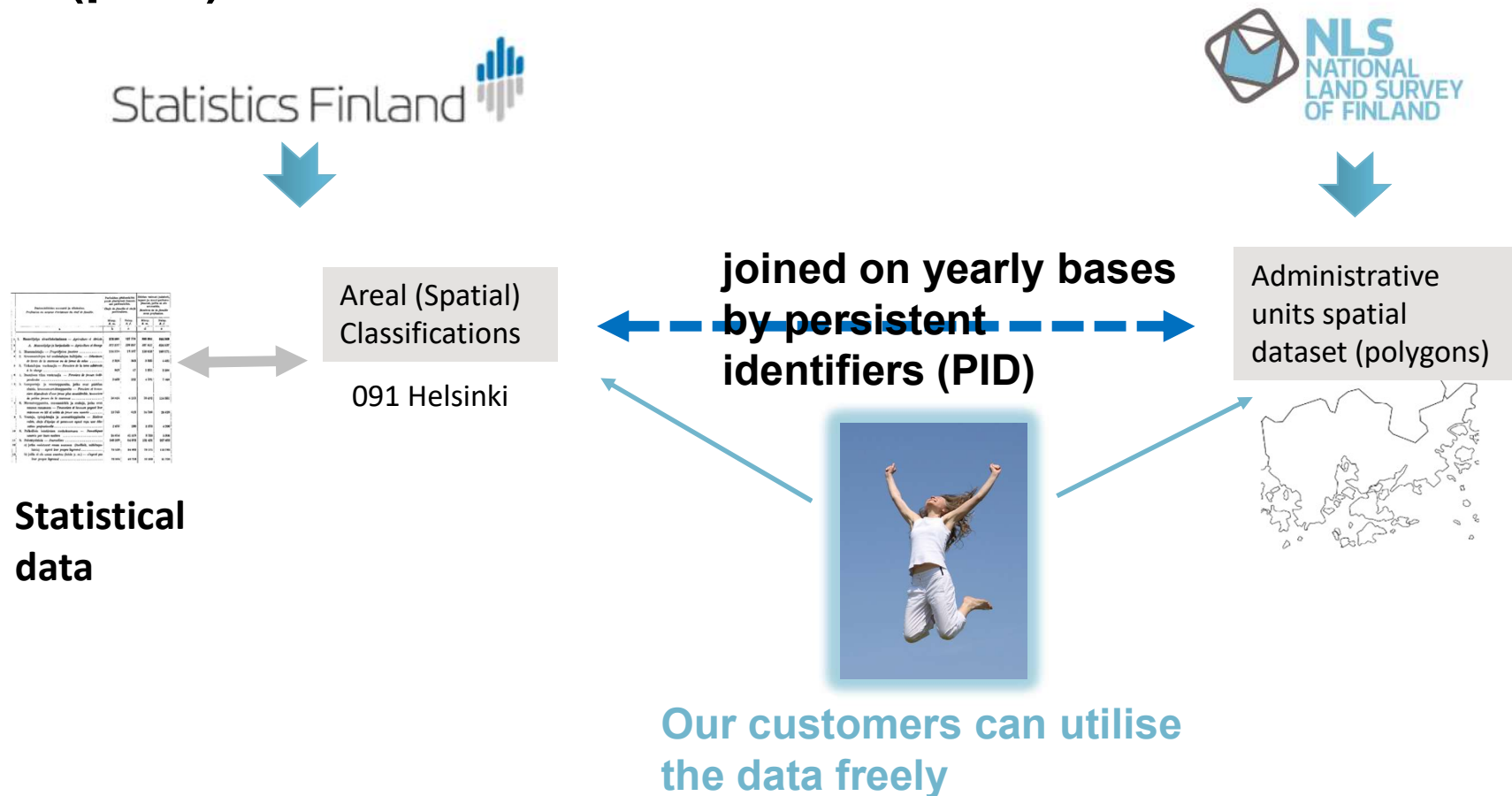
**Issue:** If a widely used thesauri like GEMET (in INSPIRE) could be enriched and transformed as ontology-form?



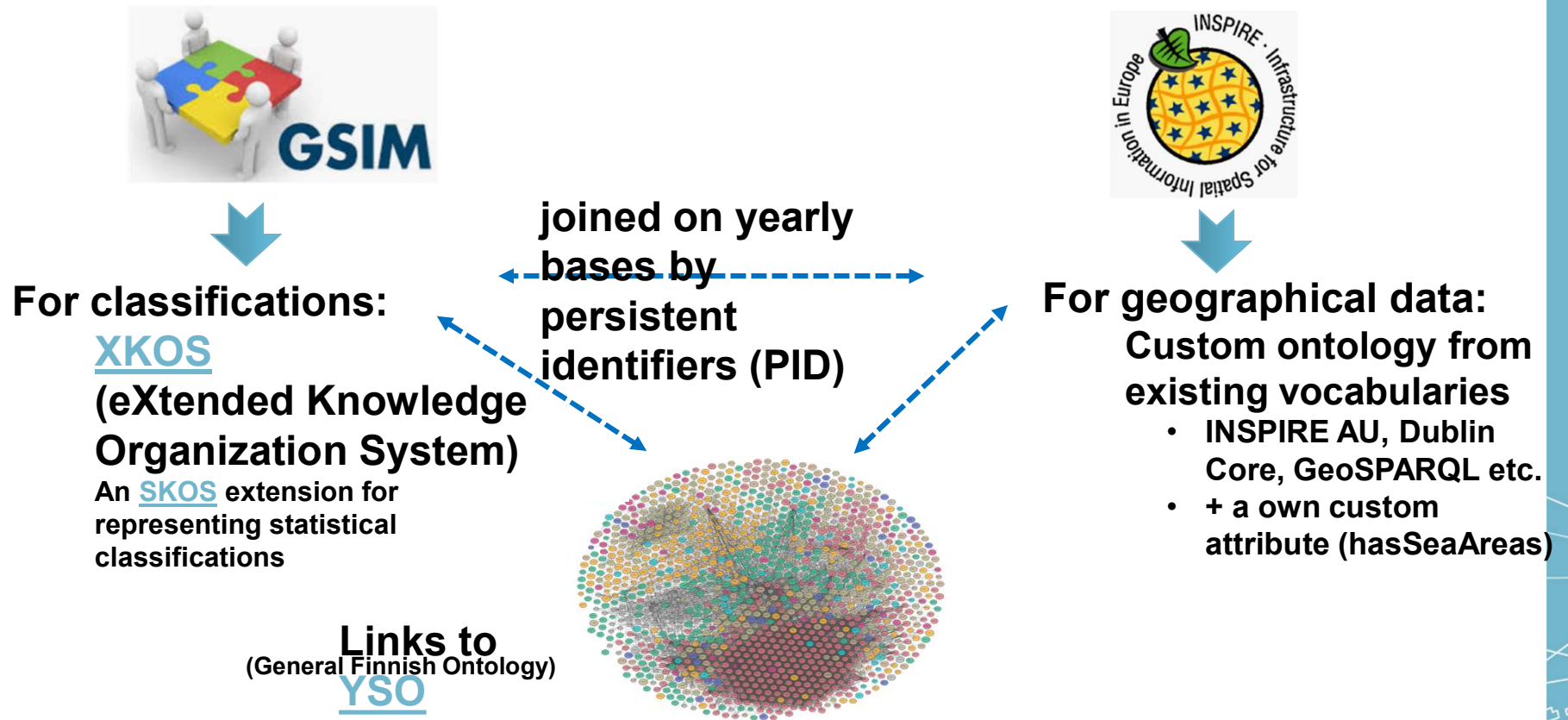
## Next: Machine readability, 'Geographic space'

- The choices:
  - Machine-readability provided through ontology **or** Elastic search indexing
- Then need for
  - Simple easy-approach guidelines for use of **SHACL, Prov-O, Shex... and their roles**
  - Improving access to metadata and understanding it (different metadata for different use)
- 'Geographic space' – revisited; some ideas
  - Collecting data from different sources (spatial and other type) on a certain area-of-interest akin 'Geographic space'
  - Present this data collection (of different views) in a graph format
    - Possibly pack it all as a hypergraph, or sub-collection-views on it as an individual graph/hypergraph?

# Providing joined data from two organisations (pilot) - IGALOD



# Standards and data models used in the solution



Picture: <https://lod-cloud.net/clouds/lod-cloud.svg>

# ALLUtion Web App visualizes the data



Areal (Spatial)  
Classifications

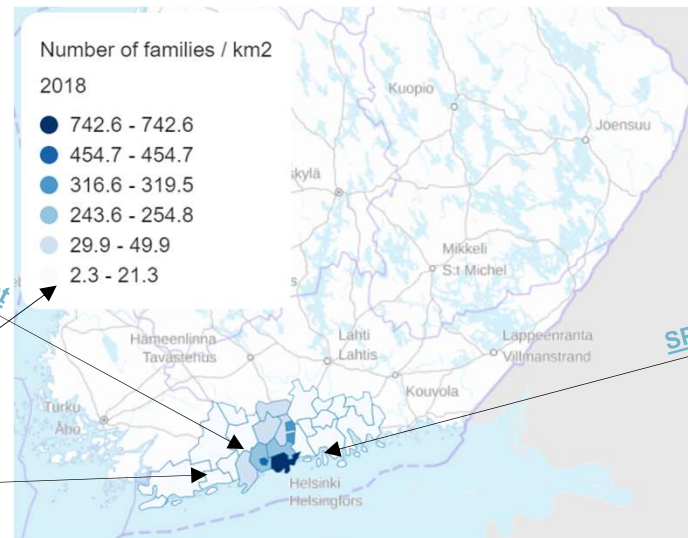
091 Helsinki

Maakuntatunnus	Maakunnan nimi	Maakunnan pinta-ala (km <sup>2</sup> )	Maakunnan väkiluku (2018)	Maakunnan keskimääräinen asutus (2018)
091	Helsinki	506,3	655,200	1,294
092	Uusimaa	4,545	1,800,000	396
093	Väestökeskittymä	1,000	100,000	100
094	Maaseutu	3,545	1,700,000	482
095	Yhteensä	5,545	2,500,000	451

Statistical  
data

SPARQL endpoint

PX-Web API



Administrative  
units spatial  
dataset

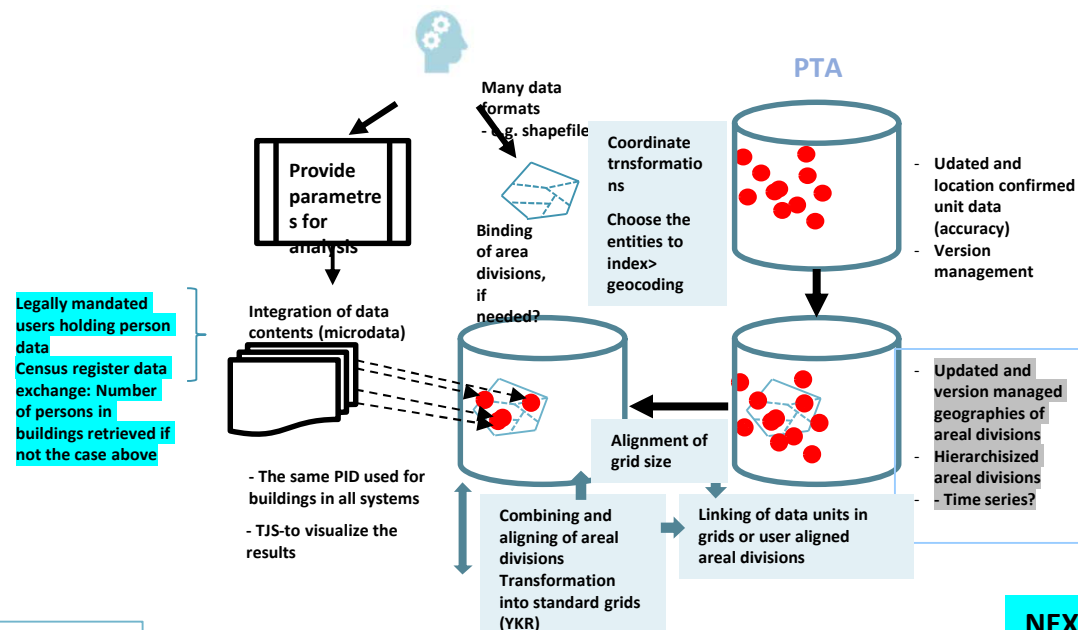


SPARQL endpoint

# Indexing service pilot – LD

In pilot phase

-for collecting data in territorial divisions (as 'Geographic spaces')



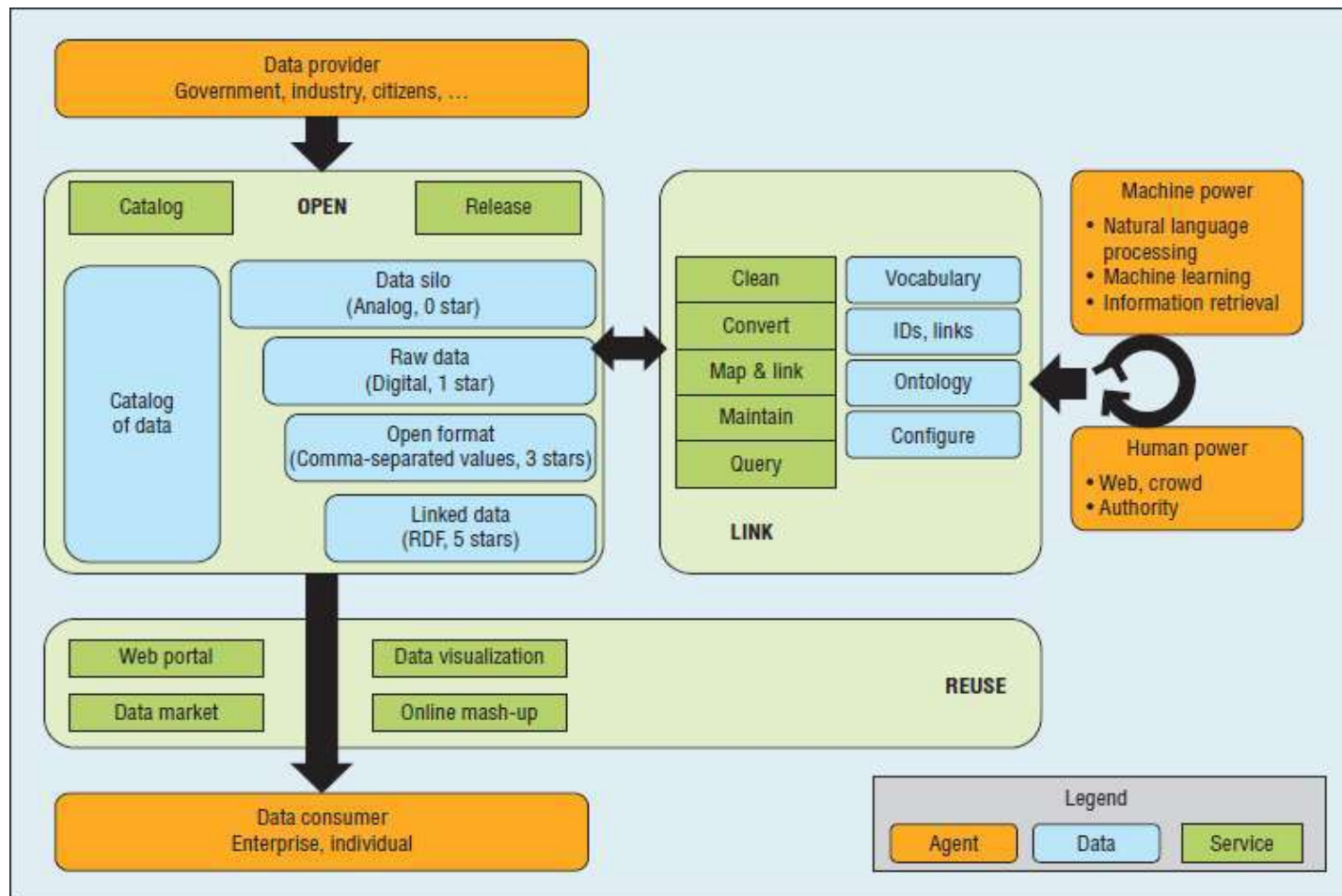
Linked data and web practices in linking geometry of territorial divisions from SDI to point-located statistical microdata (individual statistical units) to upload microdata units in those areas, and furthermore even in user-specified areas.

## SERVICE INTEGRATIONS IGALOD

IGALOD is a linked data method reusing geometries of administrative and other territorial divisions from SDI to statistic, co-operated by NLS Finland and Statistics Finland

# Ecosystem of linked [open] data

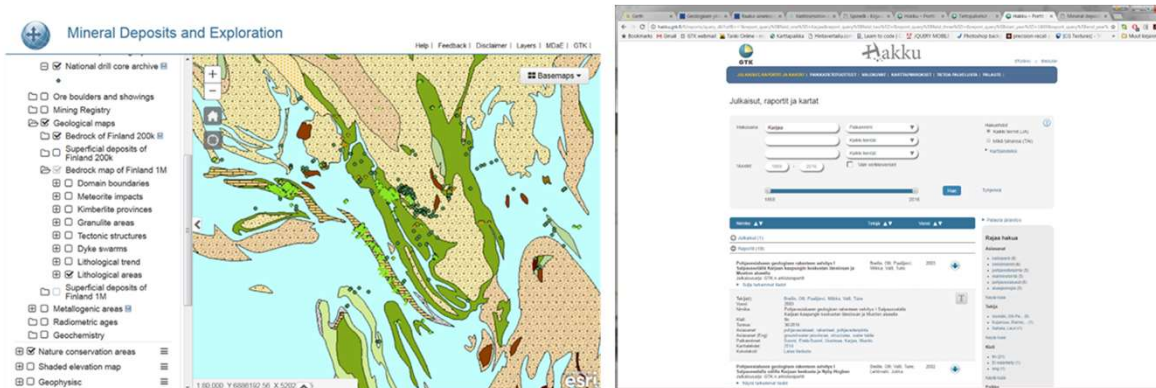
14





## Georef -

placenames for linking data – data integration of scientific report contents/data (Geology as example)



Place names bridging information & data assets

Example: Need to merge individual research reports and survey maps (Geological survey) and further across disciplines

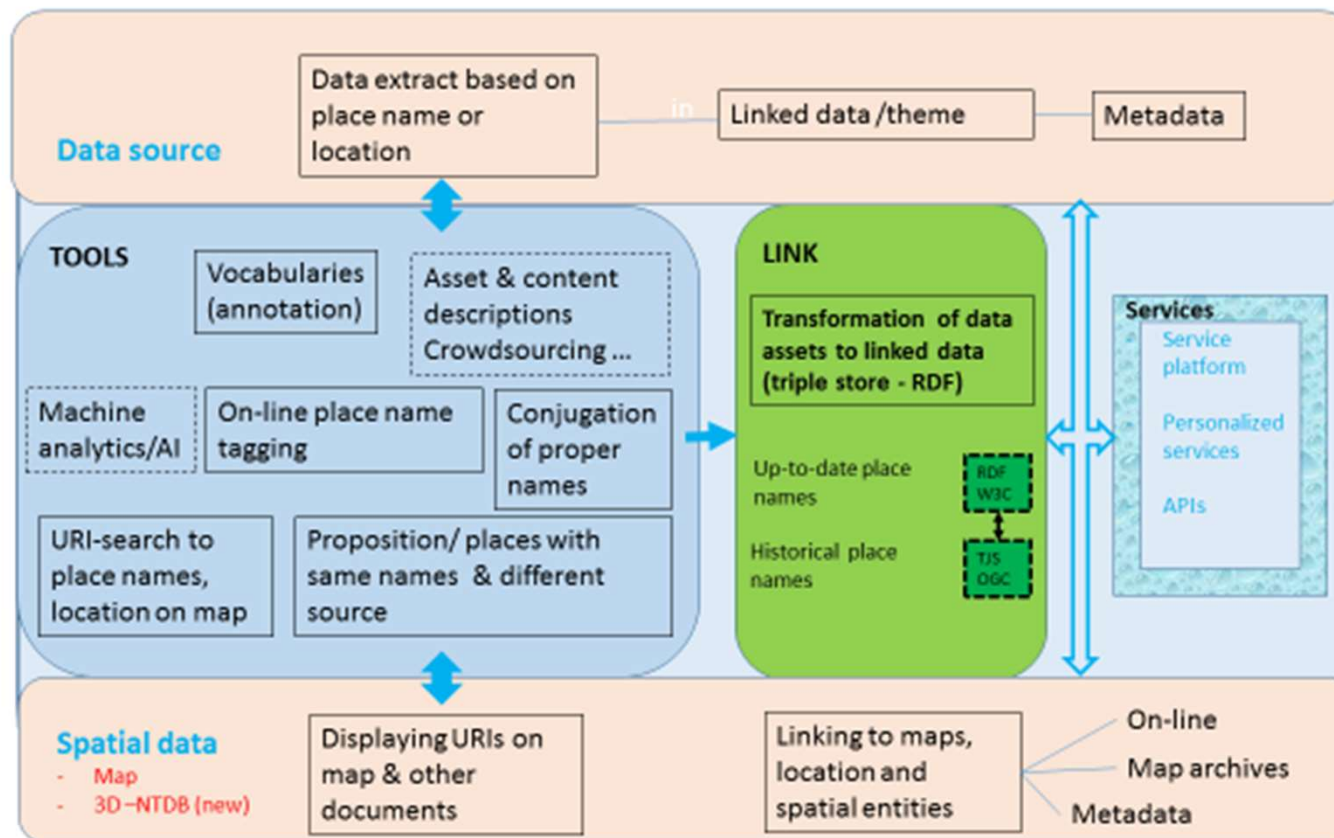


J.J. Sederholm, Pohja, Hermansö 1906

Scientific reports include place names

- To be tagged with URI's
- that are bounded by coordinates,

# Georef-linked data service and development platform





# Georef in short= geocoding by place names

Georef is targeting to enable and improve data combinations of spatial data and any other data, scoping to

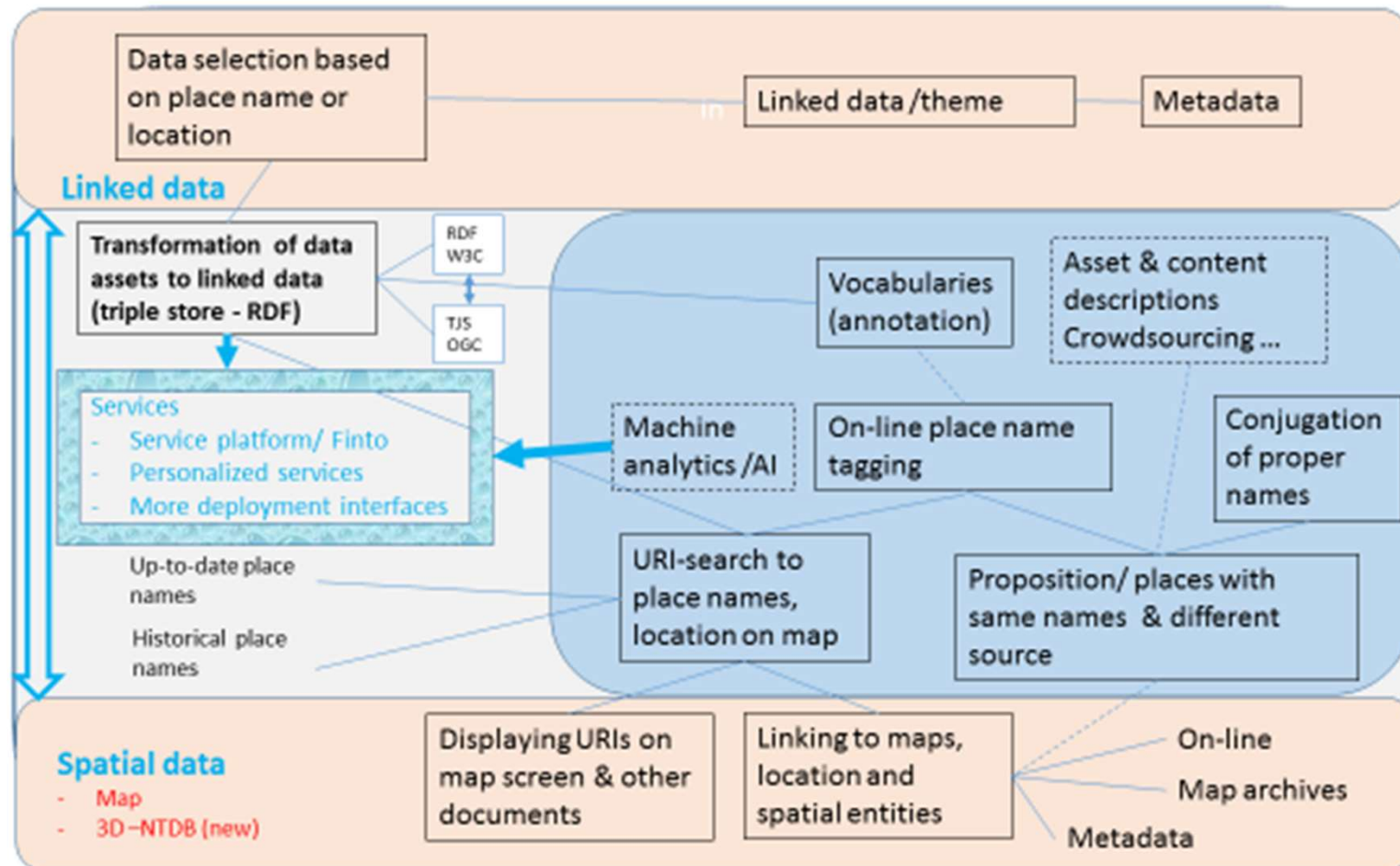
- Any web contents!
- Public services, re-use of scientific and research data across disciplines
- Re-use in media (local-national; YLE, the national broadcasting company)
- Personalized services (health, education, specialized training, immigrants...)
- Commercial services
- To link with **all** Finnish publications through finto.fi (the national ontology service)
- Crowdsourcing: Inclusion of community members in municipal decision making by providing sufficient information base and by crowdsourcing and updating contents related to places and areas to provide and create novel viewpoints and information in city development (citizen science)

Message on implementation: Local is global!

# Place names for bridging information & data assets

- A lot of information carry place names **but**
- Most information do not carry direct location data
- This applies to data assets of any format
  - scientific research reports, different types of documents; textual, images, photos, movies, music etc
- To link or combine spatial data with these other data types we employ place names binded to coordinates
- But place names are tricky too
  - Many places have the same names (Paris, Texas)
  - Endonymes, eksonymes, different conjugations etc
- To be usable they need unique identifiers i.e. httpURIs

## Georef – service and development platform; overview



Initiative in co-operation of several national research institutes, memory organisations, universities, large cities, private companies and YLE, the Finnish national broadcasting company

## Initiative on a national Land cover profile for UN sub-indicator 15.3.1 – A reference base for national SDG analysis platform

- A single SDG indicator is showing only one aspect on development
- A more comprehensive understanding is provided by illustrating how several indicators are interdependent and other territorial conditions – need for a co-op platform
- A challenge to geospatial visualization and analysis with multiple indicators and other environmental changes.
  - SDG 15.3.1 Land cover as reference data
- Data integration instead of possible siloing of single indicators
  - **Geocubes, data linking, CSW...** – *analysis tools to develop!*
  - Finland as a country with long south-north extent will have remarkable territorial differences in climate change impacts.

# Justification of the topic for UNGGIM

## Europe

### 1 (2)

- Is the topic relevant for UNGGIM Europe?
  - Basically Linked data provides a direct, dynamic and expandable method for data integration instead of often complicated and manifold, if not impossible, data transformations between stable data models and formats
  - Improves evidence based policy
- Why should the UNGGIM working through on data integration tackle this topic/task?
  - Extensive potentiality: Improved integration efficiency and timeliness; flexibility in change demands or variations through the time; resilience as regards new data sources or technical paradigms; versatile visualization or portrayal capabilities - to name a few – a disruptive information management method!

# More information

Esa Tiainen

[esa.tiainen@nls.fi](mailto:esa.tiainen@nls.fi)

<http://www.fgi.fi/fgi/research/departmentsdi-services>

[www.nls.fi](http://www.nls.fi)



**NLS**  
FINNISH GEOSPATIAL  
RESEARCH INSTITUTE  
FGI



# Implementation of linking - URI's to real world entities

- Spatial objects in NTDB model concrete objects as for reference
  - feasible for linking
  - represent real world entities as placeholders for URI
- Spatial objects-URIs of NTDBtopo and NTDBurban are linked by /id/-URI of real world entity assigned for the generalized object in NTDBtopo
- NTDB creates and updates references in RDF database (also the URI of city system)
  - If RDF database will be generated later the references can be saved in /doc/-type URI or even database.

**Real world object**  
Lake Saimaa

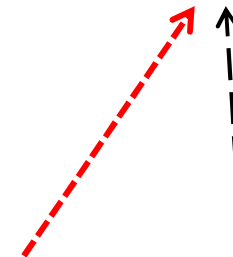


*As a spatial object*



New national TDB generates so-URI:  
<http://paikkatiedot.fi/so/100700/abc123>  
and the /id/-URI

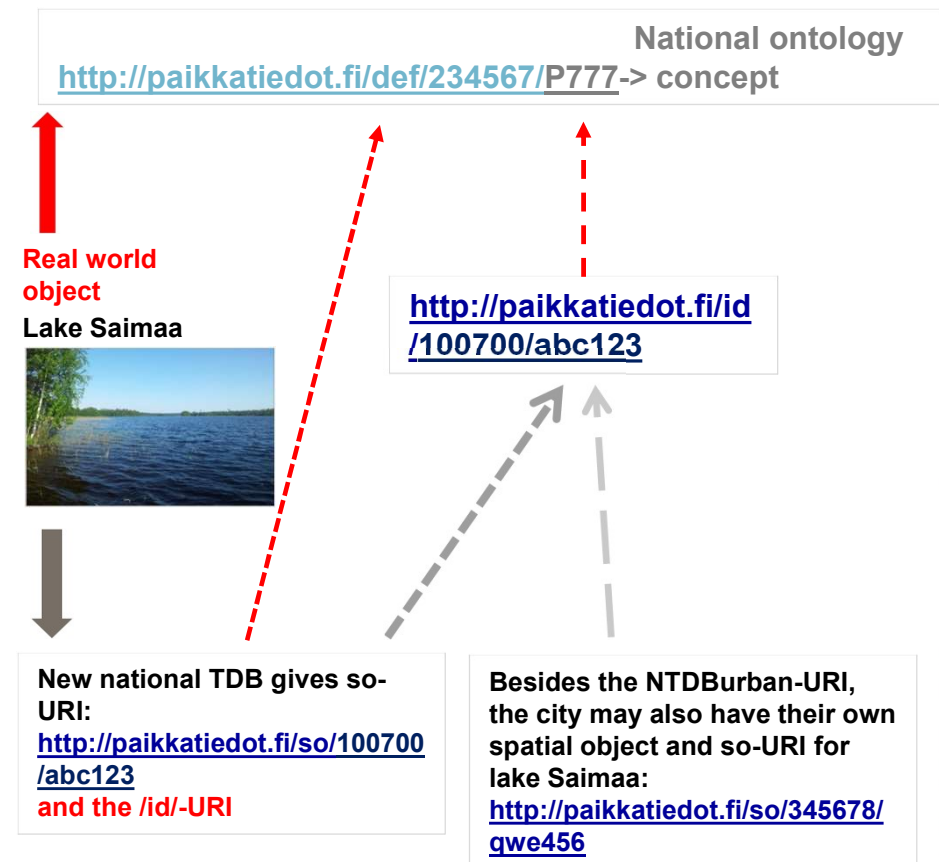
<http://paikkatiedot.fi/id/100700/abc123>



Besides the NTDBurban-URI, the city may have their own so-URI for lake Saimaa:  
<http://paikkatiedot.fi/so/345678/qwe456>

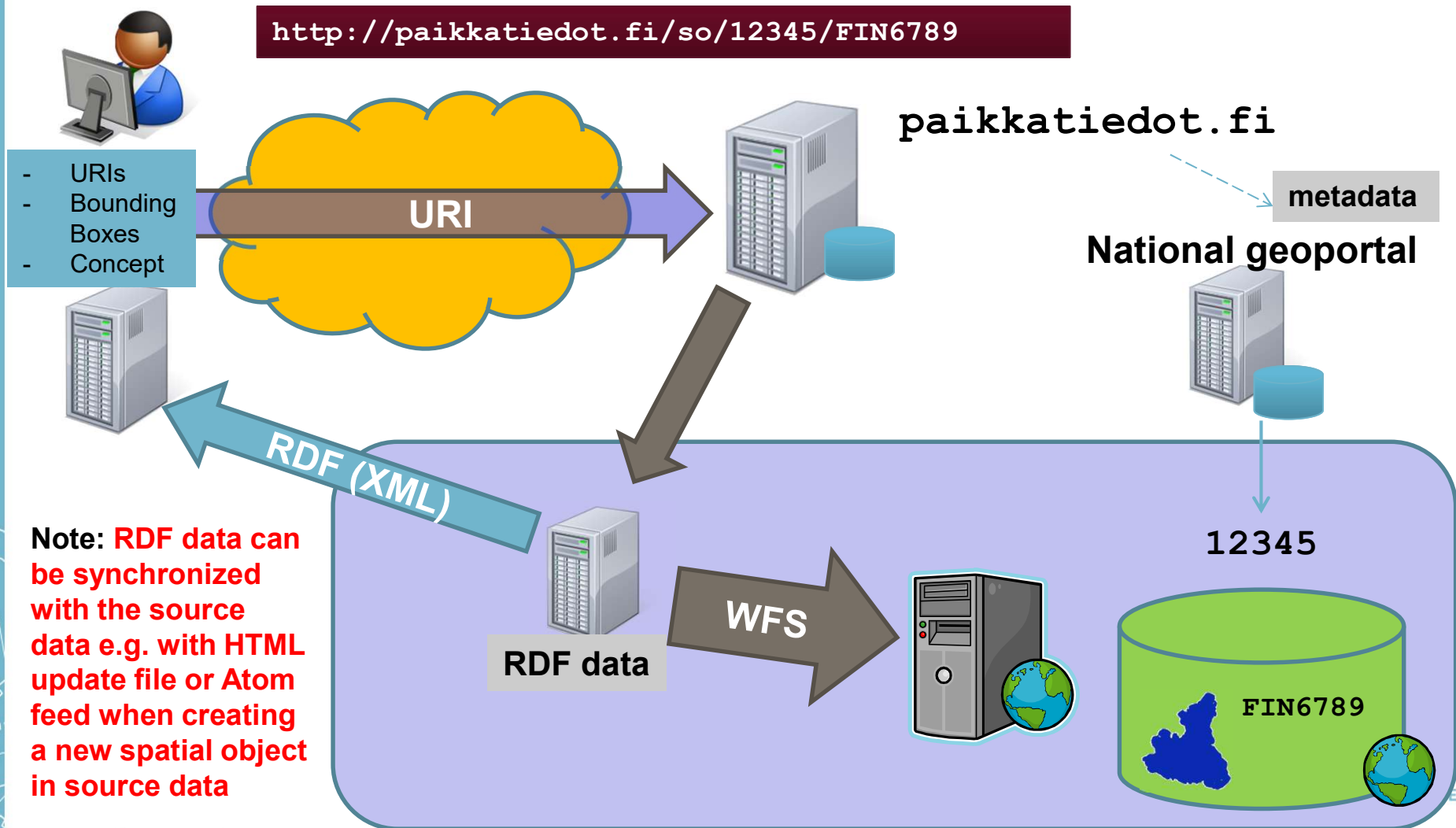
# Implementation of linking - Concepts

- Linking to concepts to integrate spatial and non-spatial data
- Linking from real world URIs (or spatial object URIs)
- And upload to RDF database
- **Not yet in the piloting agenda**
- Special case: Place names
  - Place name service platform
  - to provide applications for tagging different assets and further linking through RDF
  - Presentation on pilot and technical solutions in EuroSDR2016 (Eero Hietanen NLS FGI)





# Technical structure



# Thank you for interest!

More information

Esa Tiainen

senior adviser / johtava asiantuntija

National Land Survey/ Maanmittauslaitos

Finnish Geospatial Research Institute FGI

[esa.tiainen@nls.fi](mailto:esa.tiainen@nls.fi)

[www.maanmittauslaitos.fi](http://www.maanmittauslaitos.fi)

[www.fgi.fi](http://www.fgi.fi)

