

# A cloud based platform for Linking and Managing Geodata

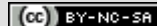
Dimitris Kotzinos

University of Cergy Pontoise and FORTH-ICS

Kyriakos Kritikos and Yannis Roussakis

FORTH-ICS

This content by [the InGeoCloudS consortium members](http://www.ingeoclouids.eu/) is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License](http://creativecommons.org/licenses/by-nc-sa/3.0/). Based on a work at <http://www.ingeoclouids.eu/>.

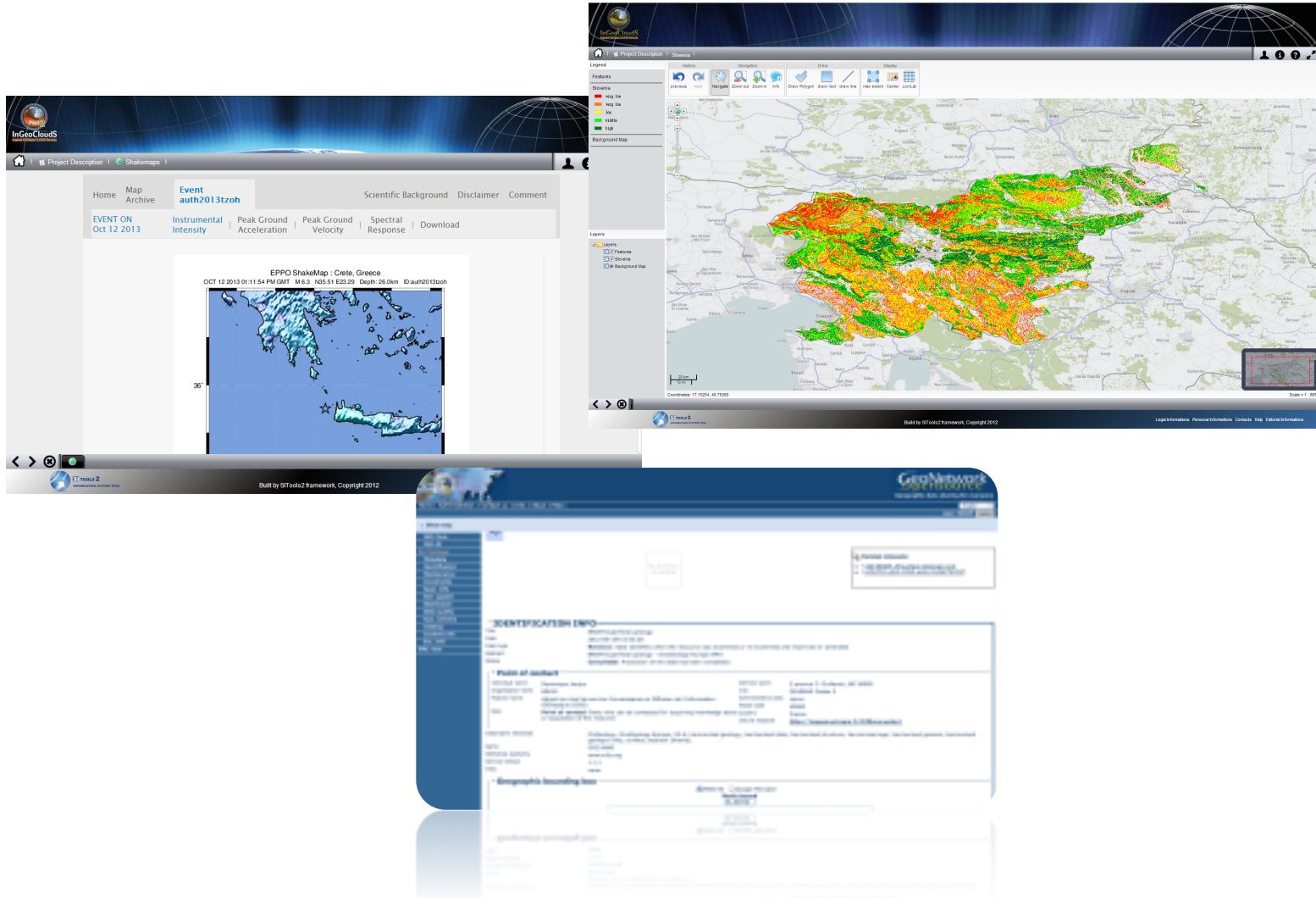


# Objectives in a nutshell

- Design and build a cloud infrastructure for public agencies in the spatial-environmental field
- Provide an innovate (flexible,...) infrastructure for geo-data services
- “Move” public services to a cloud-based infrastructure.
- Integrate geo-data by exploiting a Linked Data “model”



# Integrated Geo Applications (www.ingeoclouds.eu)



# Linked (Open) Data as a Service

## Publishing Linked Data

- ▶ URI construction
- ▶ Conceptual Model
- ▶ Storage as RDF files or SPARQL endpoints

## Querying Linked Data

- ▶ SPARQL
- ▶ GeoSPARQL

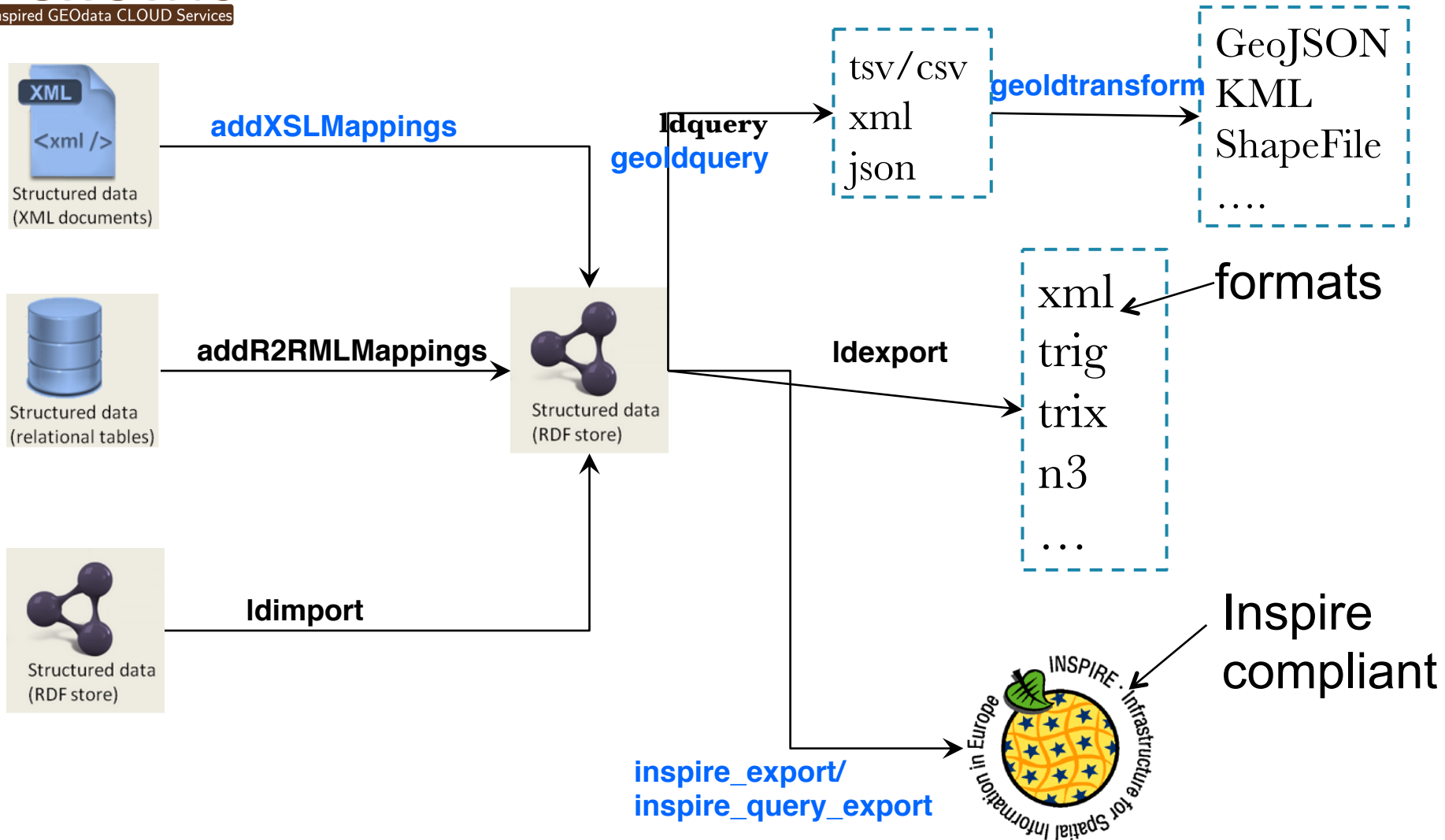
## Updating Linked Data

- ▶ SPARUL
- ▶ Synchronization with original sources

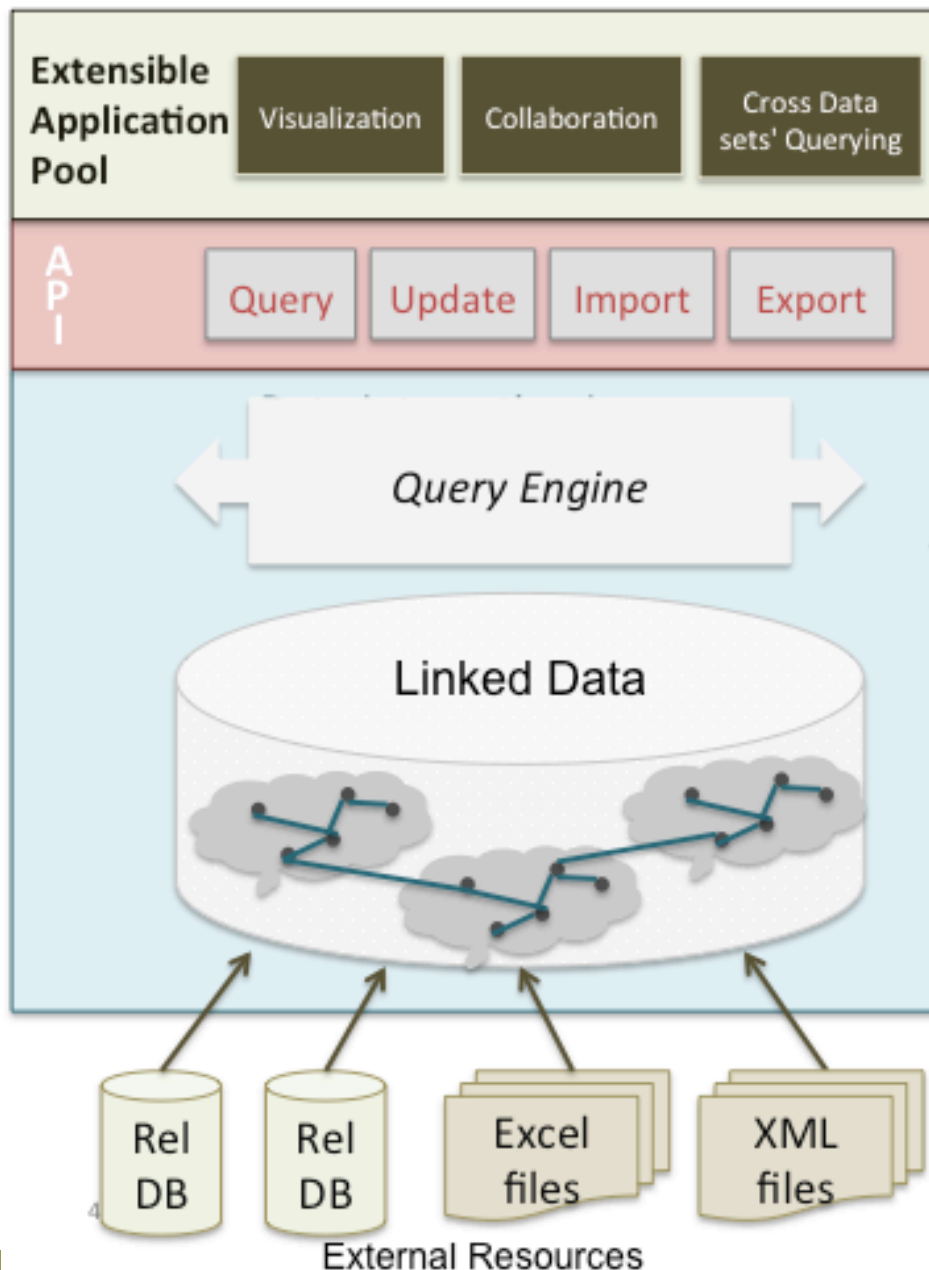


**InGeoCloudS**  
Inspired GEOdata CLOUD Services

# Linked Data Services



# Linked (Open) Data as a Service



Abstraction layer for data access

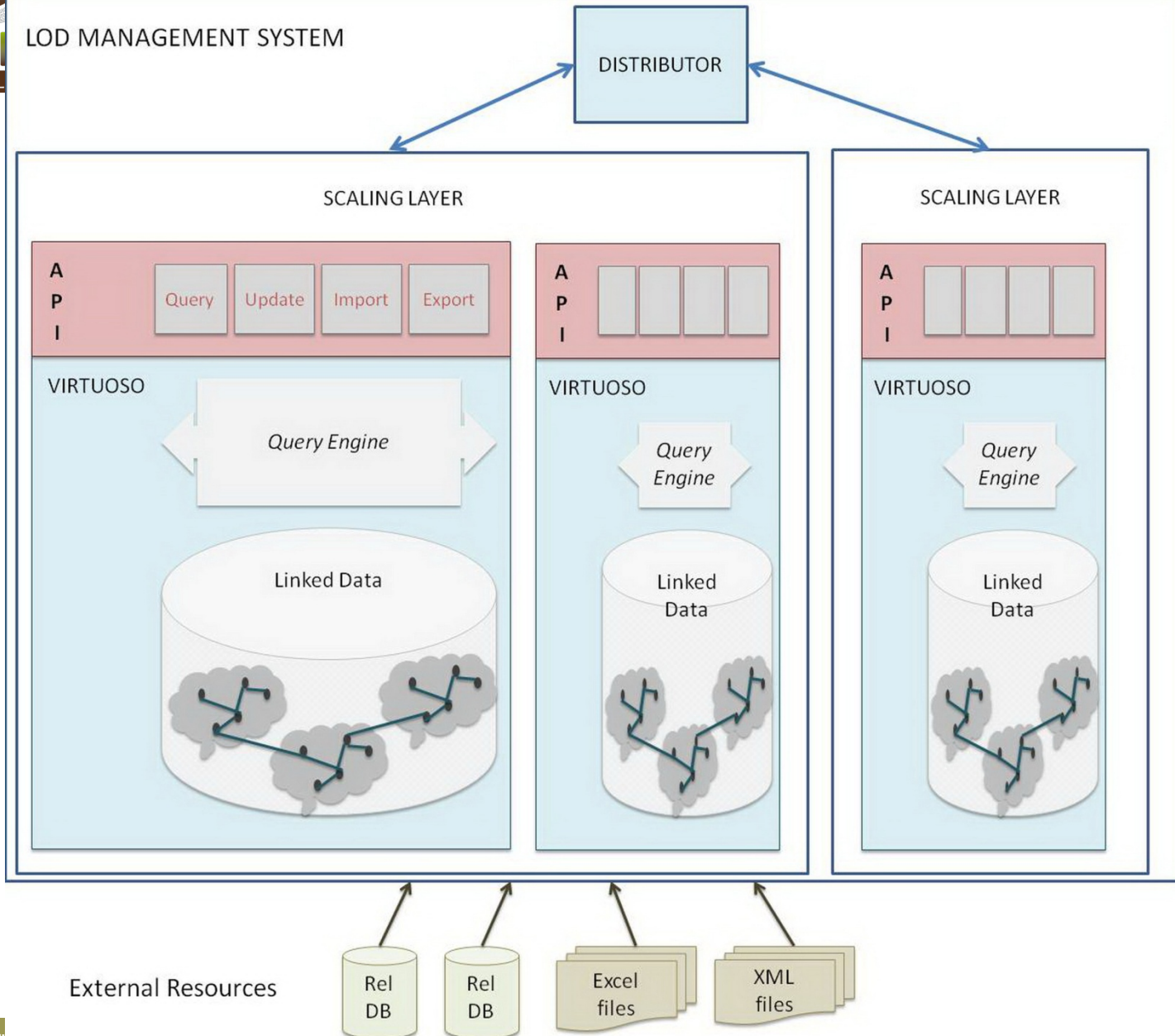
*abstract the applications from the specific setup of the data management service (such as local vs. remote, federation, and distribution)*

Beyond Data Access

- Enabling automation of discovery, composition, and use of datasets
- Data Markets
- Online Visualization Services
- Data Publishing Solutions
- Data Aggregators
- BI / Analytics as a Service



LOD MANAGEMENT SYSTEM





# LOD Integration & Publishing (I)

## URI Creation

(semi) automatic creation of URIs if needed

Two alternatives for publishing LOD:

1. Create and import RDF-based descriptions of data-sets via particular LMS method

- ▶ Data update process must be controlled by performing SPARUL updates via particular LMS method
- ▶ Data provider responsibility to keep synchronized relational & RDF data
  - ▶ A perfect synchronization may be also not required as it may incur costs -> second alternative becomes more preferable

# LOD Integration & Publishing (II)

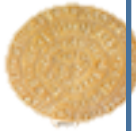
2. Data provider publishes relational data of his/her data sets + provides a mapping file in R2RML to enable the synchronization of relational to RDF data (by executing LMS method)
  - ▶ System takes care of this synchronization
  - ▶ Relational storage in the way used many years + additional RDF storage for the data with automatic one-way synchronization between the two
  - ▶ Provider should have a good knowledge of RDF

## R2RML:

- ▶ W3C recommendation since 2012
- ▶ Can specify customized mappings between RDB & RDF data
- ▶ R2RML specification is just a RDF graph in Turtle
- ▶ No specific implementation is imposed

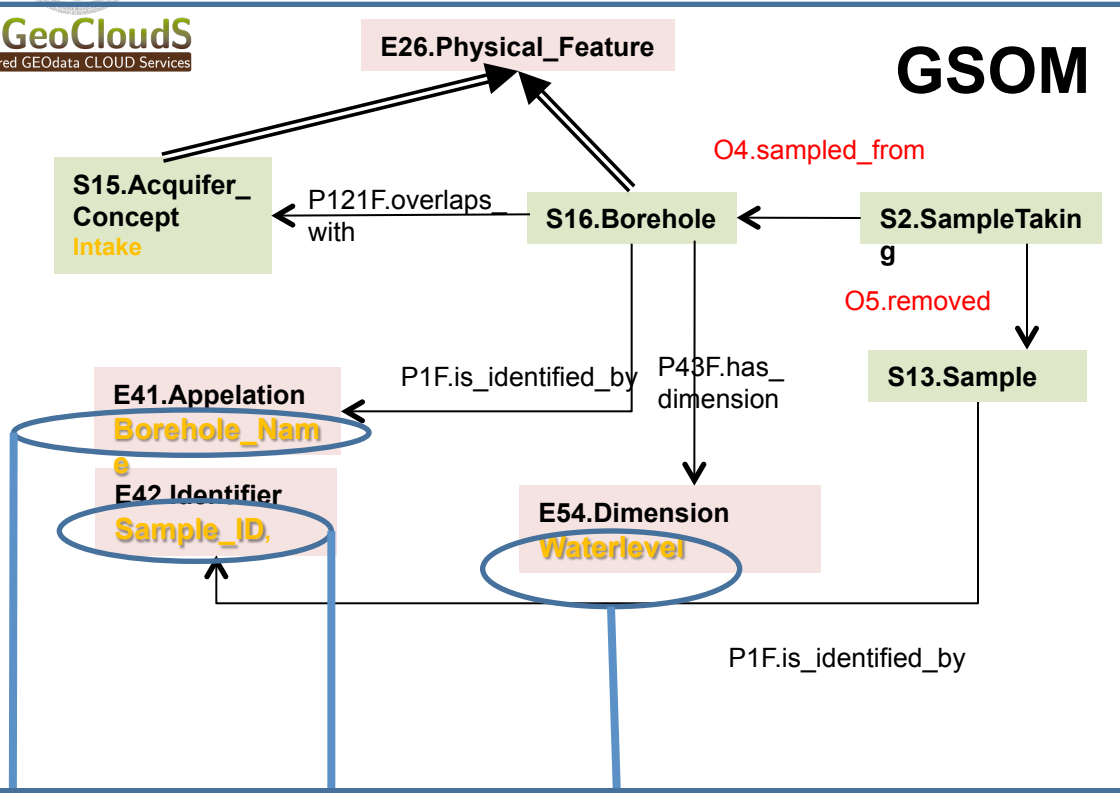


# R2RML



InGeoCloudS  
Inspired GEOdata CLOUD Services

## GSOM



**URI Identification:**  
<http://orgURL/SampleID/XYZ>

<b>Borehole Name</b>	<b>Sample ID</b>	<b>WaterLevel</b>
B1	XYZ	Level1

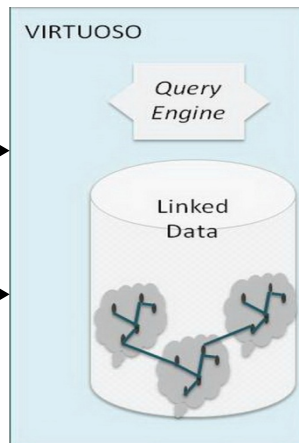
## Borehole Relational Model



**RDB**

**Publication**

**Synchronization**

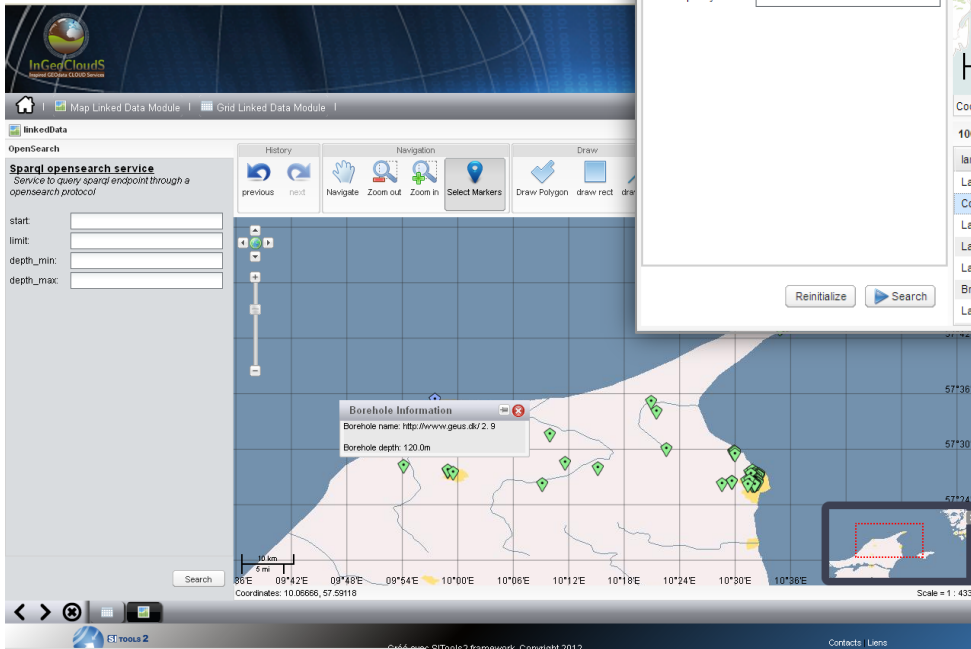


# Export Linked Data into INSPIRE

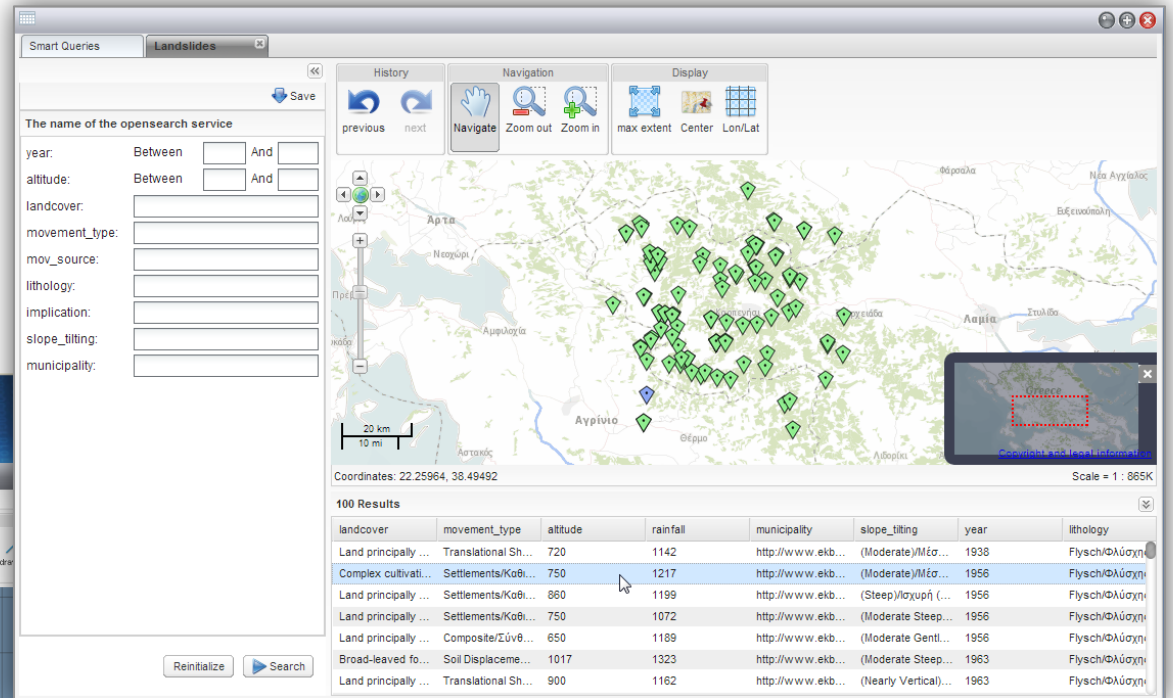


- We created a our model(GSOM)-to-INSPIRE mapping mechanism which allows to transform GSOM-based RDF data, acquired from SPARQL queries, into XML-based INSPIRE-compliant data
- Exported Data are made INSPIRE-compliant
- However INSPIRE has limitations which do not allow for a complete export of all the GSOM notions (i.e., we cannot express all the GSOM notions in INSPIRE notions)
- Ability to export:
  - Directly the results of the queries
  - Whole INSPIRE themes

# Querying Linked Data Store



The screenshot shows the InGeoCloudS web interface. On the left, there is a search panel with fields for 'start', 'limit', 'depth\_min', and 'depth\_max'. The main area displays a map of Greece with several green diamond markers. A popup window titled 'Borehole Information' is open, showing details for a borehole: 'Borehole name: http://www.geus.dk/2.9' and 'Borehole depth: 120.0m'. The interface includes navigation and display controls at the top.



The screenshot shows the InGeoCloudS web interface with a search query for 'Landslides'. The search panel on the left has the following fields filled in:

- year: Between [ ] And [ ]
- altitude: Between [ ] And [ ]
- landcover: [ ]
- movement\_type: [ ]
- mov\_source: [ ]
- lithology: [ ]
- implication: [ ]
- slope\_tilting: [ ]
- municipality: [ ]

The map on the right shows a large number of green diamond markers representing landslide locations. Below the map, the coordinates are 22.25964, 38.49492 and the scale is 1:865K. A table below the map displays 100 results:

landcover	movement_type	altitude	rainfall	municipality	slope_tilting	year	lithology
Land princially ...	Translational Sh...	720	1142	http://www.ekb...	(Moderate)/Μέσ...	1938	Flysch/Φλύσση
Complex cultivati...	Settlements/Koθ...	750	1217	http://www.ekb...	(Moderate)/Μέσ...	1956	Flysch/Φλύσση
Land princially ...	Settlements/Koθ...	860	1199	http://www.ekb...	(Steep)/Ισχυρή (...	1956	Flysch/Φλύσση
Land princially ...	Settlements/Koθ...	750	1072	http://www.ekb...	(Moderate Steep...	1956	Flysch/Φλύσση
Land princially ...	Composite/Ζύθ...	650	1189	http://www.ekb...	(Moderate Gentl...	1956	Flysch/Φλύσση
Broad-leaved fo...	Soil Displaceme...	1017	1323	http://www.ekb...	(Moderate Steep...	1963	Flysch/Φλύσση
Land princially ...	Translational Sh...	900	1162	http://www.ekb...	(Nearly Vertical)...	1963	Flysch/Φλύσση

# GeoProcessing

Geoprocessing implemented as a WPS service:  
refers to ordinary **kriging** interpolation

WPS is provisioned in the ElasticWebServer component  
and uses parameters:

- Given by the user
- Fetched from InGeoCloudS Triplestore using LD-API





# Linking Data Services

- Combine external and internal data sources
- Linking data to each other
- Retrieve information on demand

**External and Internal Data Integration**

Sort:

Results Format:

Min Latitude:

Max Latitude:

Min Longitude:

Max Longitude:

Min Date:

Max Date:

Landslide Type:

Moving Source:

Feature Dim. Type:

Feature Dim. Value:

Feature Dim. Monotonicity:

Landslide Dim. Type:

Landslide Dim. Value:

Landslide Dim. Monotonicity:

Amount of Results:

[Data Sources](#)





# Challenges related to Linked Data and the Cloud - “Political” Challenges

**InGeoCloudS**  
Inspired GEOdata CLOUD Services

- User adoption
  - Users are happy with what they have and they patiently wait until the problem surfaces
  - Users with weak or no infrastructure are most receptive to turn to Cloud and/or Linked Data paradigms than others
  - User see the cloud as a platform with more resources (memory, storage, processing) but they still want their applications to run there unchanged
  - **Security, privacy and trust**
- Public vs. Private Clouds
- Publicly owned vs. Company-owned Clouds



# Challenges related to Linked Data and the Cloud - Scientific Challenges

**InGeoCloudS**  
Inspired GEOdata CLOUD Services

- (Big) Linked Data storage
  - No “native” RDF triplestore that exploits the Cloud capabilities
  - Data security / privacy / control also a technical issue, especially when we deal with big data that you cannot monitor precisely
- Querying (Big) Linked Data
  - Fast querying for big data (many times the queries are simple but the volume is big)
  - Support for geospatial queries
  - (Fast) Indexing for (RDF) geospatial data
- “Exchanging” (Big) Linked Data
  - Data providers want to keep their own infrastructures and synchronize data between them and the cloud! Issues on synchronization and exchange of big data ...

# Conclusions

Proposed a scalable, geo-spatial LOD as-a-Service management system deployed on Amazon cloud

- ▶ Distributes query load + scales-up/down when CPU utilization surpasses specific thresholds
- ▶ Exposes REST-based service with LOD management methods
- ▶ Provides two different ways for publishing open geo-spatial data sets
- ▶ Provides various export possibilities

# Thank you and Questions

[www.ingecLOUDS.eu](http://www.ingecLOUDS.eu)

[Dimitrios.KotzinOS@u-cergy.fr](mailto:Dimitrios.KotzinOS@u-cergy.fr)

