Representing and Querying Linked Geospatial Data

Kostis Kyzirakos
kostis@cwi.nl

Centrum voor Wiskunde en Informatica
Database Architectures group
Amsterdam
The Netherlands

Geonovum
April 11, 2014

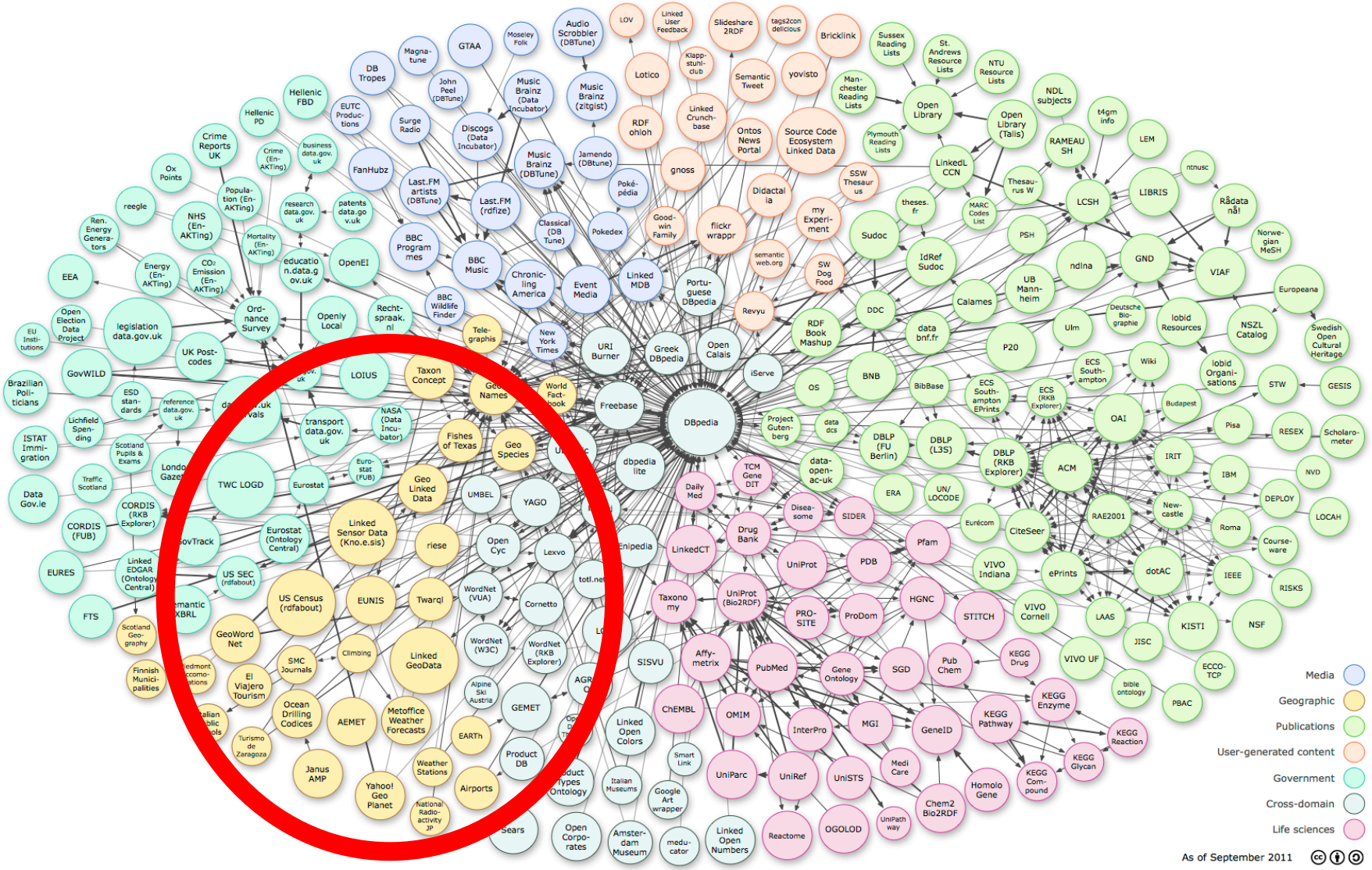
Outline

- The data model **stRDF** and the query language **stSPARQL**
- The system **Strabon**
- Visualizing time-evolving geometries using **Sextant**
- Real-Time Fire Monitoring application
- Conclusions

- The data model stRDF and the query language stSPARQL
- The system Strabon
- Visualizing time-evolving geometries using Sextant
- Real-Time Fire Monitoring application
- Conclusions

The data model stRDF and the query language stSPARQL

Linked Open Data Cloud



RDF: Resource Description Framework

W3C recommendation

RDF is a **graph data model**

- **Resources** are described in terms of properties and property values using RDF **statements**
- Statements are represented as **triples**, consisting of a **subject**, **predicate** and **object**.

The Data Model stRDF

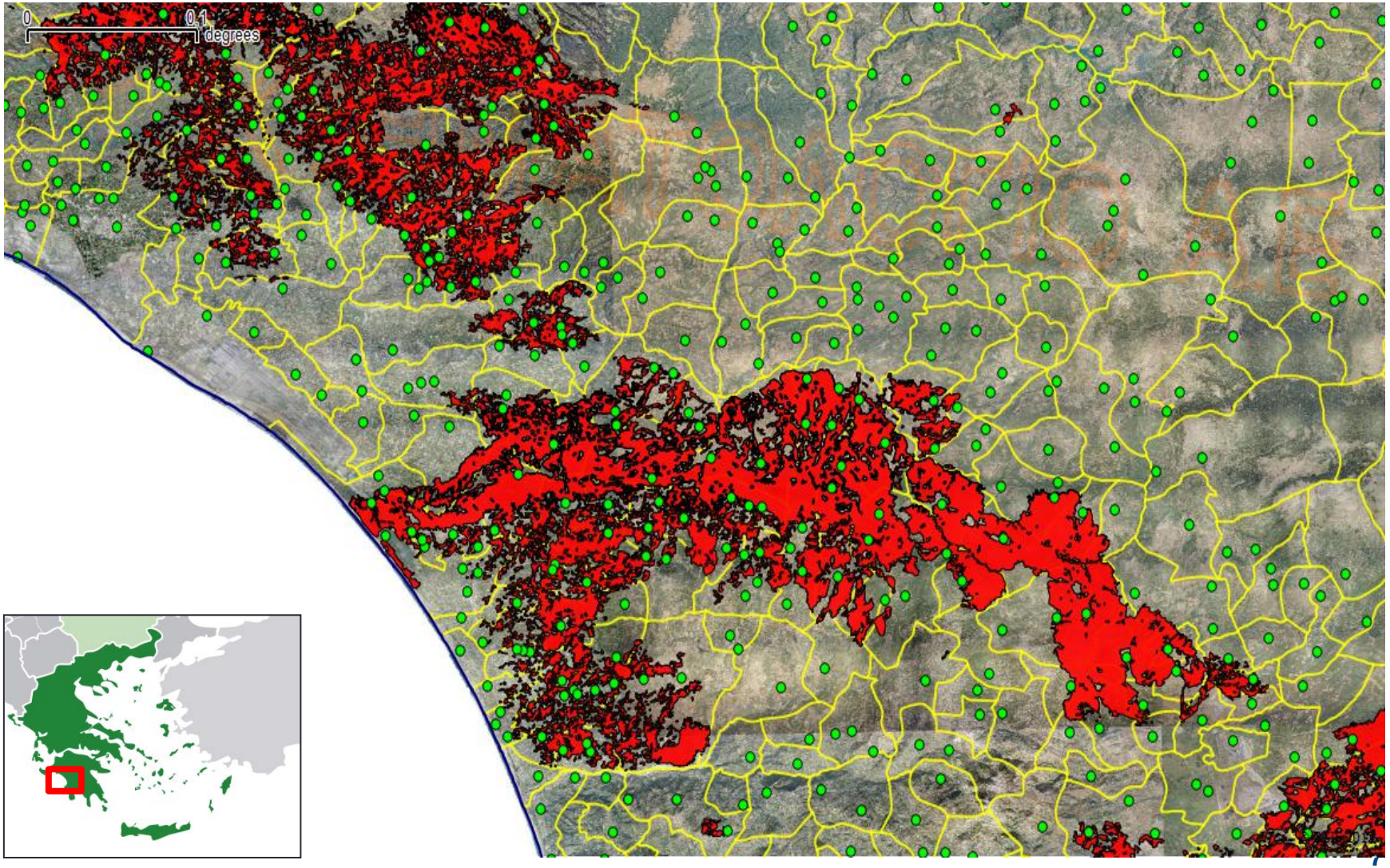
- **stRDF** stands for **spatial / temporal RDF.**
- It is an extension of the W3C standard RDF for the representation of **geospatial data that may change over time**.

[ESWC 2010, ISWC 2012]

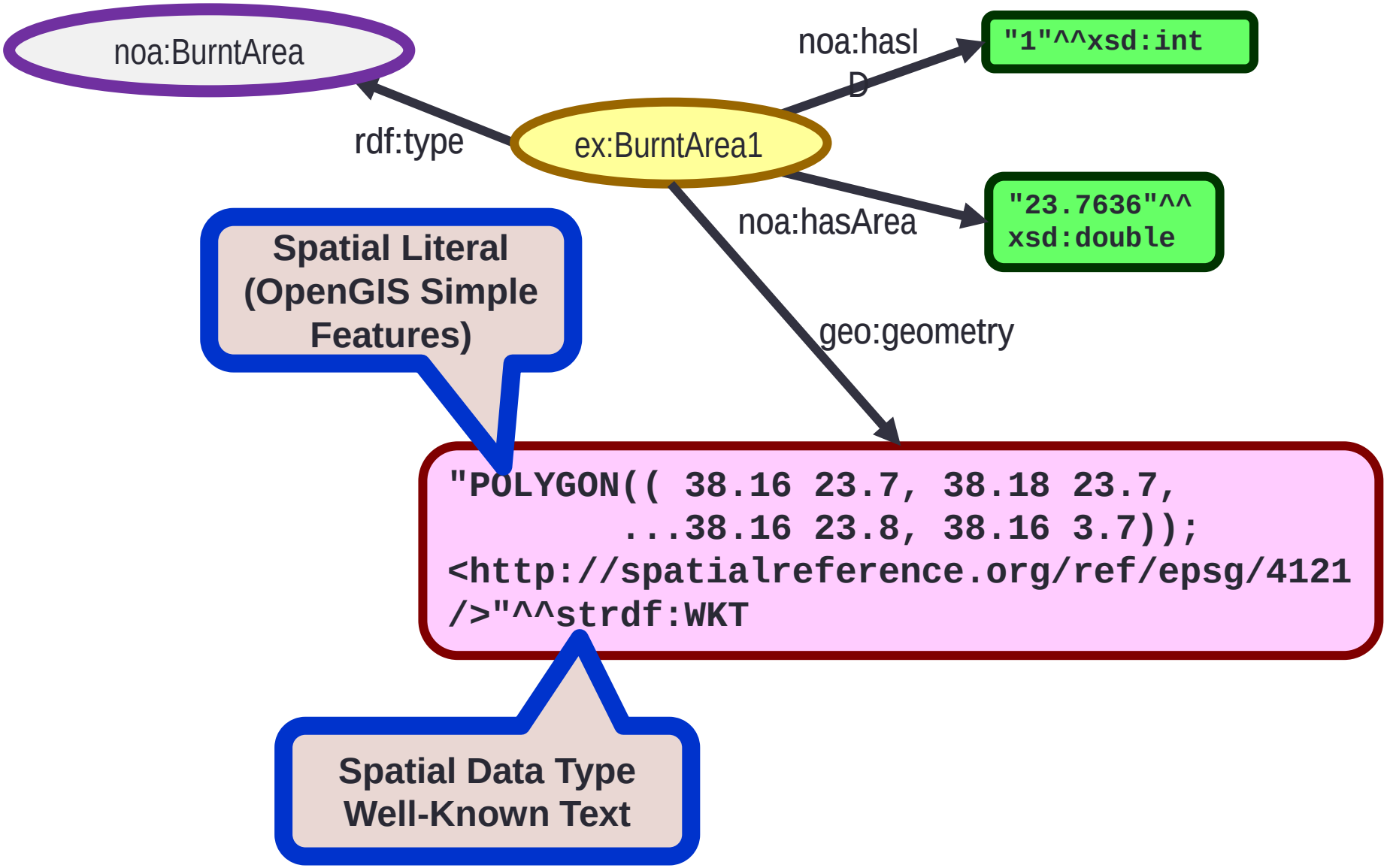
- stRDF extends RDF with:
  - **Spatial literals** encoded in OGC standards Well-Known Text or GML
    - **New datatypes** for spatial literals (`strdf:WKT`, `strdf:GML` and `strdf:geometry`)
  - **Temporal literals** can be either periods or instants
    - **New datatype** for temporal literals (`strdf:period`)
    - Placed as the fourth component of a triple to denote valid time

[ESWC 2010 ESWC 2013]

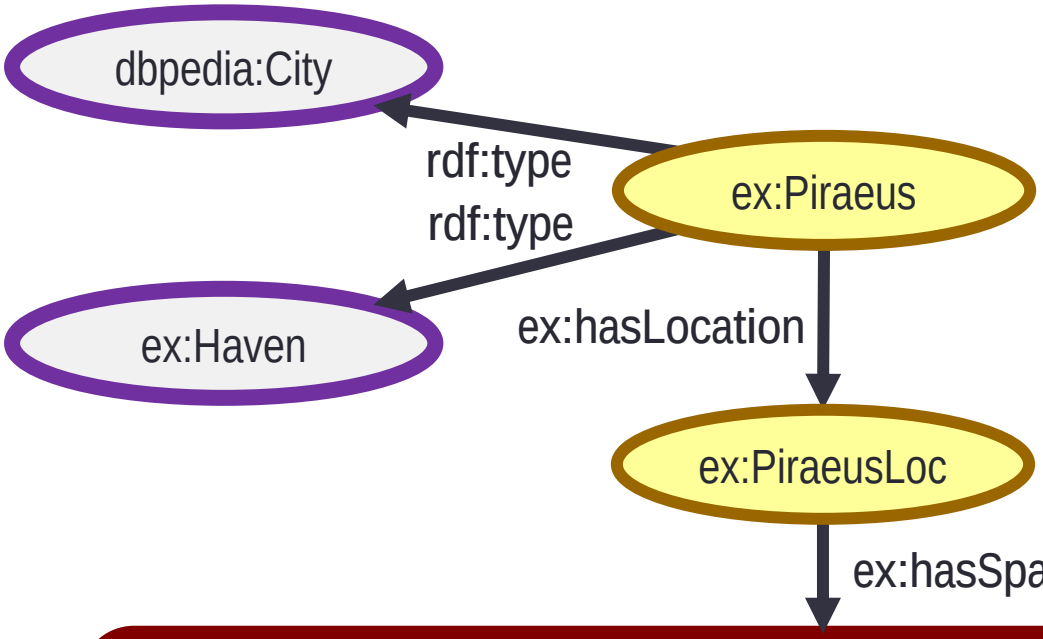
RDF: An example



stRDF: An example (1/2)



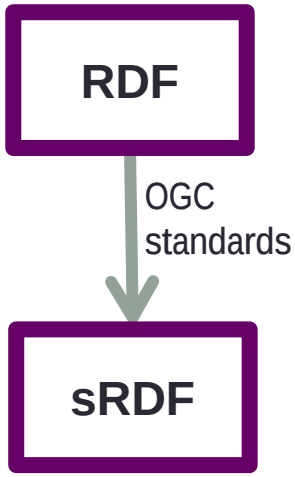
stRDF: An example

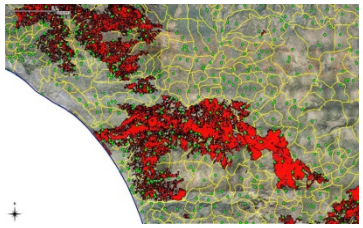



```
"POLYGON((38.16 23.7, 38.18 23.7,
      ... 38.16 23.8, 38.16 3.7));
<http://spatialreference.org/ref/epsg/4121/>
"^^strdf:WKT
```

Spatial Literal (OpenGIS Simple Features)

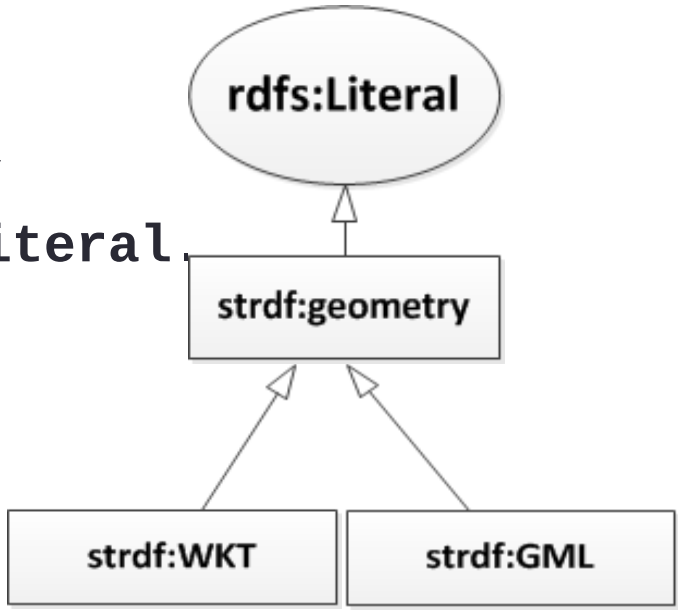
Spatial Data Type Well-Known Text





The stRDF Data Model

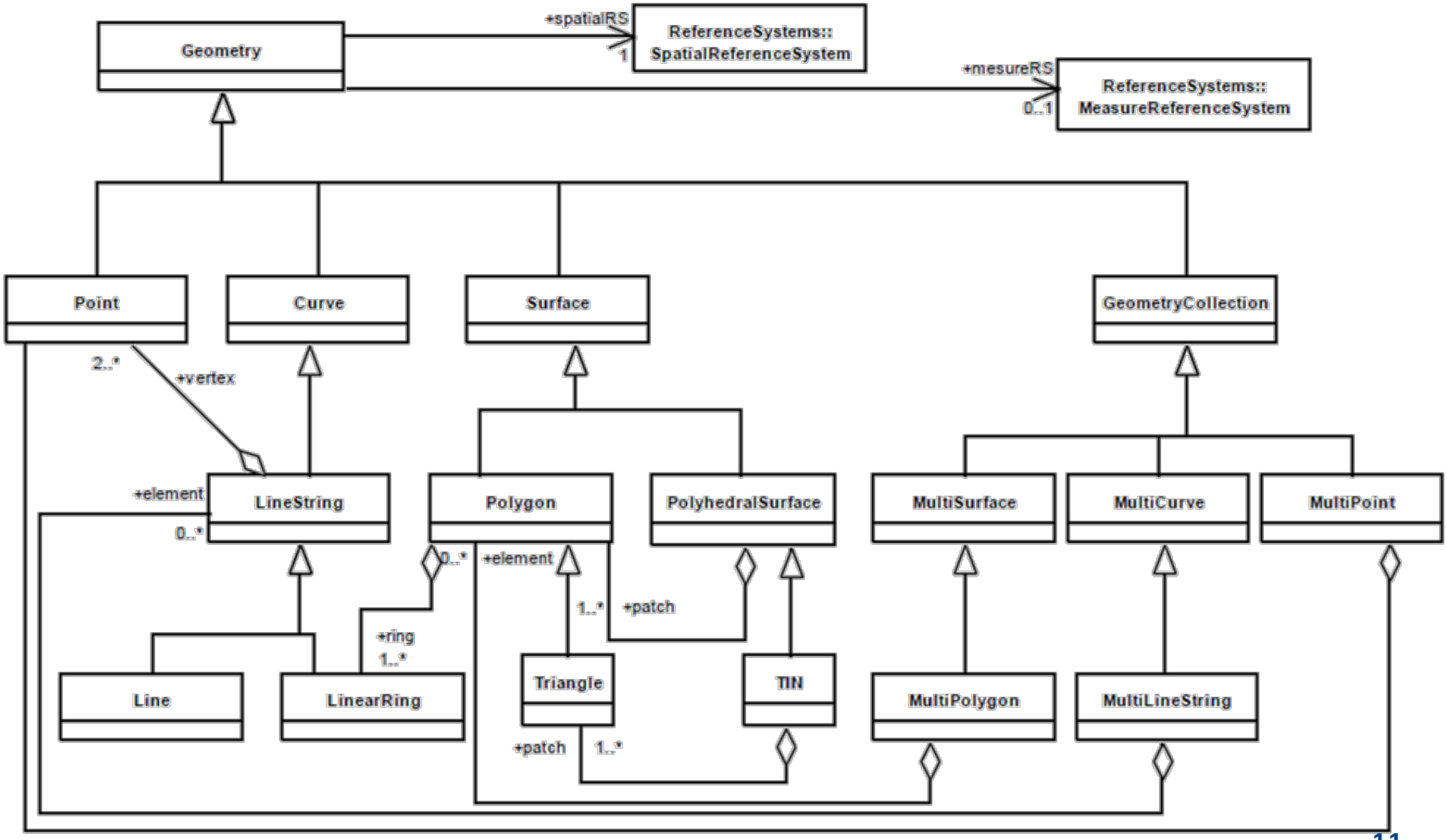
```
strdf:geometry rdf:type rdfs:Datatype;
               rdfs:subClassOf rdfs:Literal.
```




```
strdf:WKT   rdf:type rdfs:Datatype;
            rdfs:subClassOf strdf:geometry.

strdf:GML   rdf:type    rdfs:Datatype;
            rdfs:subClassOf   strdf:geometry.
```

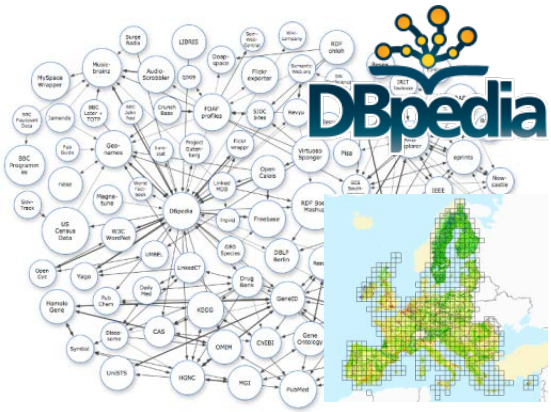
WKT Class Hierarchy



stSPARQL: An example (1/2)

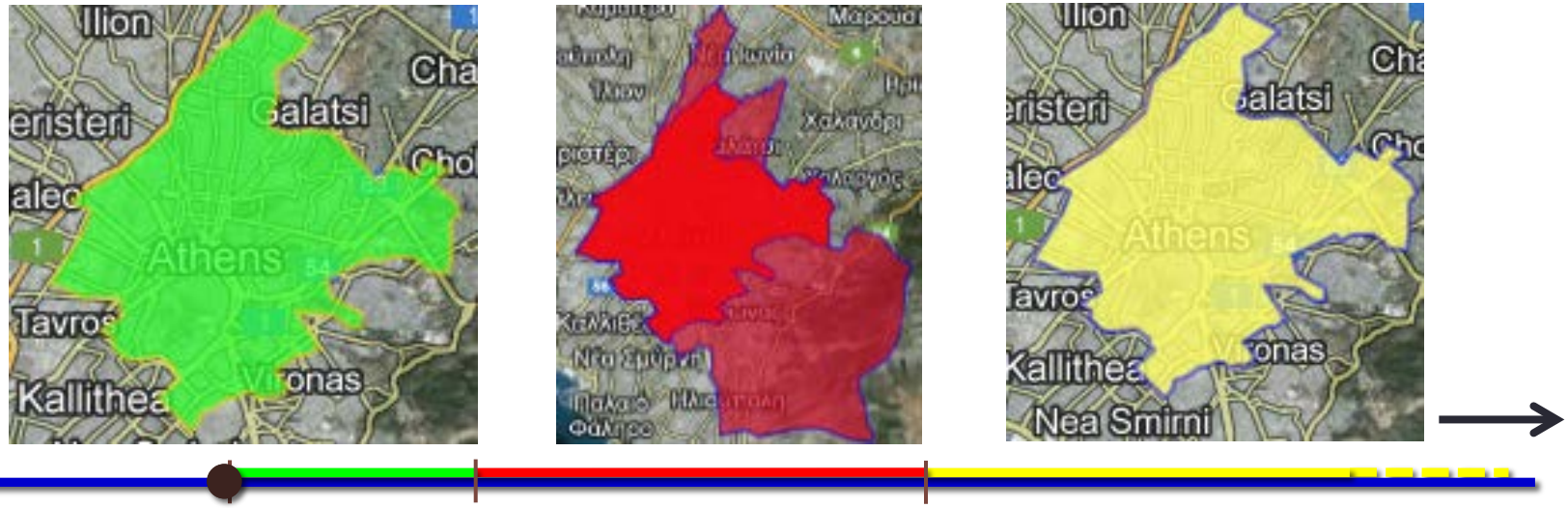
- Find all burned forests within 10kms of a city

```
select ?BA ?BAGEO
where { ?Rrdf:type noa:Region ;
        geo:geometry ?RGEO ;
      noa:hasCorineLandCoverUse ?F .
?F    rdfs:subClassOf clc:Forests .
?CITY    rdf:type dbpedia:City
         geo:geometry ?CGEO .
?BA   rdf:type noa:BurntArea ;
         geo:geometry ?BAGEO .
filter( strdf:Intersect(?RGEO,?BAGEO) &&
      strdf:Distance(?BAGEO,?CGEO,uom:km)<10)
```



Spatial Functions (OGC Simple Feature Access)

stRDF: An example (2/2)




```
clc:region1 clc:hasLandCover clc:Forest .
       "[2006-08-   25T11:00:00+02,2007-08-
25T11:00:00+02)"^^strdf:period .

noa:ba1 rdf:type noa:BurntArea
       "[2007-08-25T11:00:00+02,2009-08-
  25T11:00:00+02)"^^strdf:period .
clc:region1 clc:hasLandCover clc:AgriculturalArea
  "[2009-08-25T11:00:00+02, "UC")"^^strdf:period .
```

stSPARQL: An example (2/2)

- Find all areas that were forests in 2006 and got burned later within 10kms of a city

```
select ?BA ?BAGEO
where { ?R rdf:type noa:Region ;
        geo:geometry ?RGEO ;
        noa:hasCorineLandCoverUse ?F ?t1 .
?F   rdfs:subClassOf clc:Forests .
?CITY   rdf:type dbpedia:City ;
        geo:geometry ?CGEO .
        rdf:type noa:BurntArea ?t2;
        geo:geometry ?BAGEO .
filter( strdf:Intersect(?RGEO,?BAGEO) &&
        strdf:Distance(?BAGEO,?CGEO,uom:km)<10)
filter( strdf:during(?t1, "[2006-01-01:00:00:01,
 2006-01-01:23:59:59]"^^strdf:period)) &&
 strdf:before(?t1, ?t2) }
```

Temporal constant and extension function

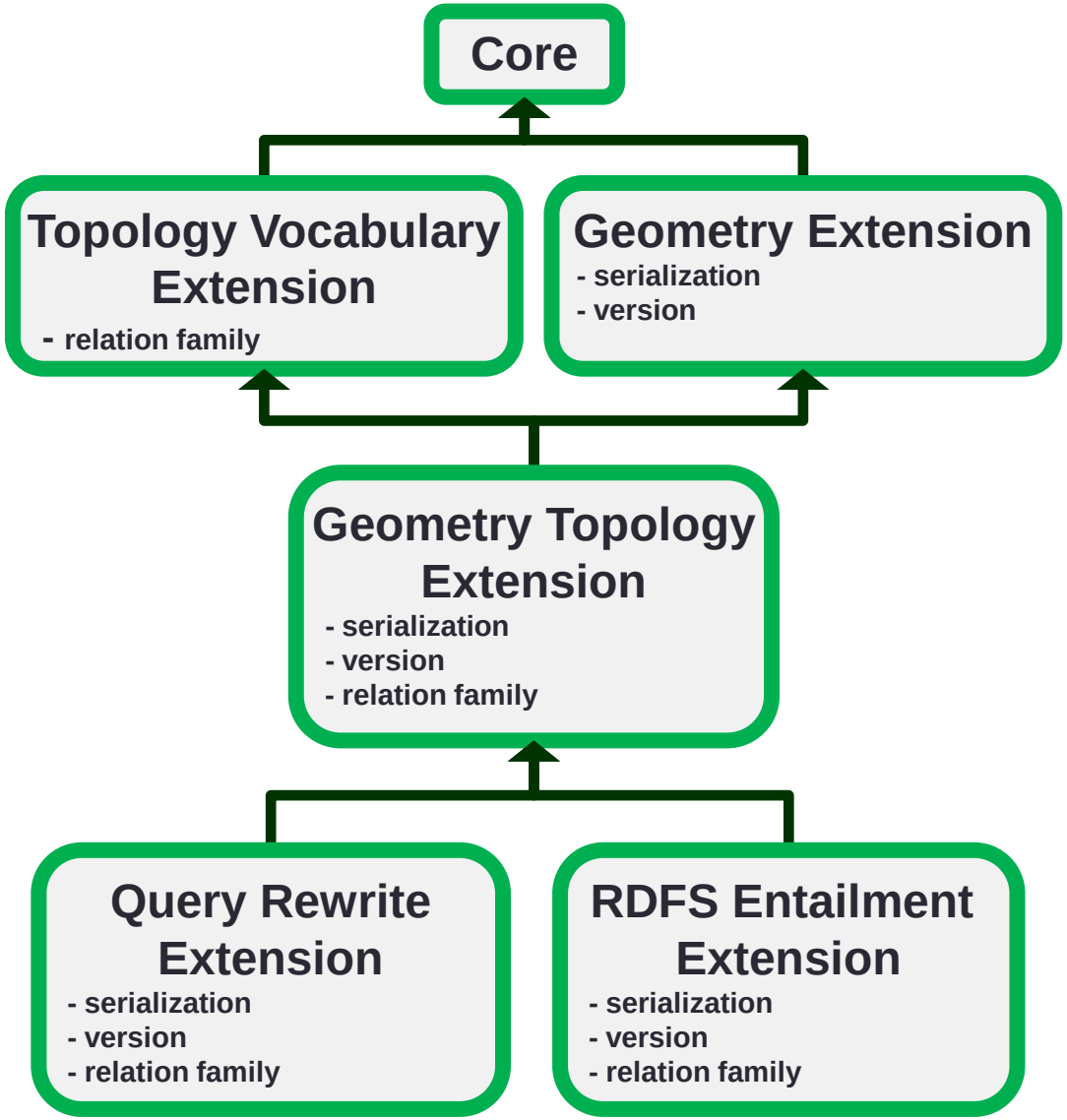
Spatial Functions (OGC Simple Feature Access)

stSPARQL: Geospatial SPARQL 1.1

We define a **SPARQL extension function** for each function defined in the **OpenGIS Simple Features Access** standard

- Basic functions
  - Get a property of a geometry (e.g., **`strdf:srid`**)
  - Get the desired representation of a geometry (e.g., **`strdf:AsText`**)
  - Test whether a certain condition holds (e.g., **`strdf:IsEmpty`**, **`strdf:IsSimple`**)
- Functions for **testing topological spatial relationships** (e.g., **`strdf:equals`**, **`strdf:intersects`**)
- **Spatial analysis** functions
  - Construct new geometric objects from existing geometric objects (e.g., **`strdf:buffer`**, **`strdf:intersection`**, **`strdf:convexHull`**)
  - Spatial metric functions (e.g., **`strdf:distance`**, **`strdf:area`**)
- **Spatial aggregate** functions (e.g., **`strdf:union`**, **`strdf:extent`**)
- We add a set of **temporal functions** (superset of Allen's functions) as SPARQL extension functions

The OGC Standard GeoSPARQL



Parameters

- **Serialization**
  - WKT
  - GML
- **Relation Family**
  - Simple Features
  - RCC-8
  - Egenhofer

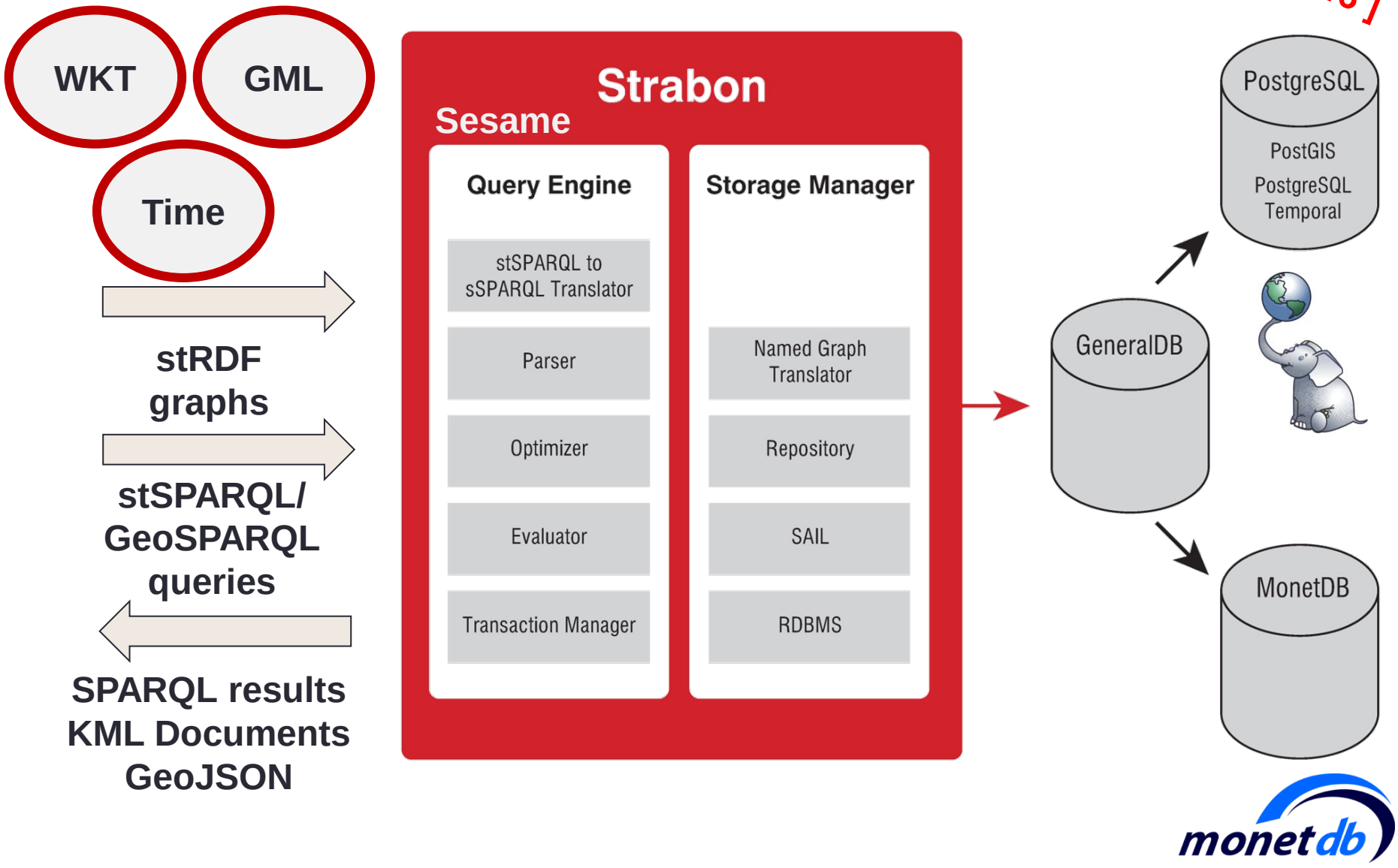
- The data model stRDF and the query language stSPARQL
- The system Strabon
- Visualizing time-evolving geometries using Sextant
- Real-Time Fire Monitoring application
- Conclusions

The system Strabon

http://strabon.di.uoa.gr

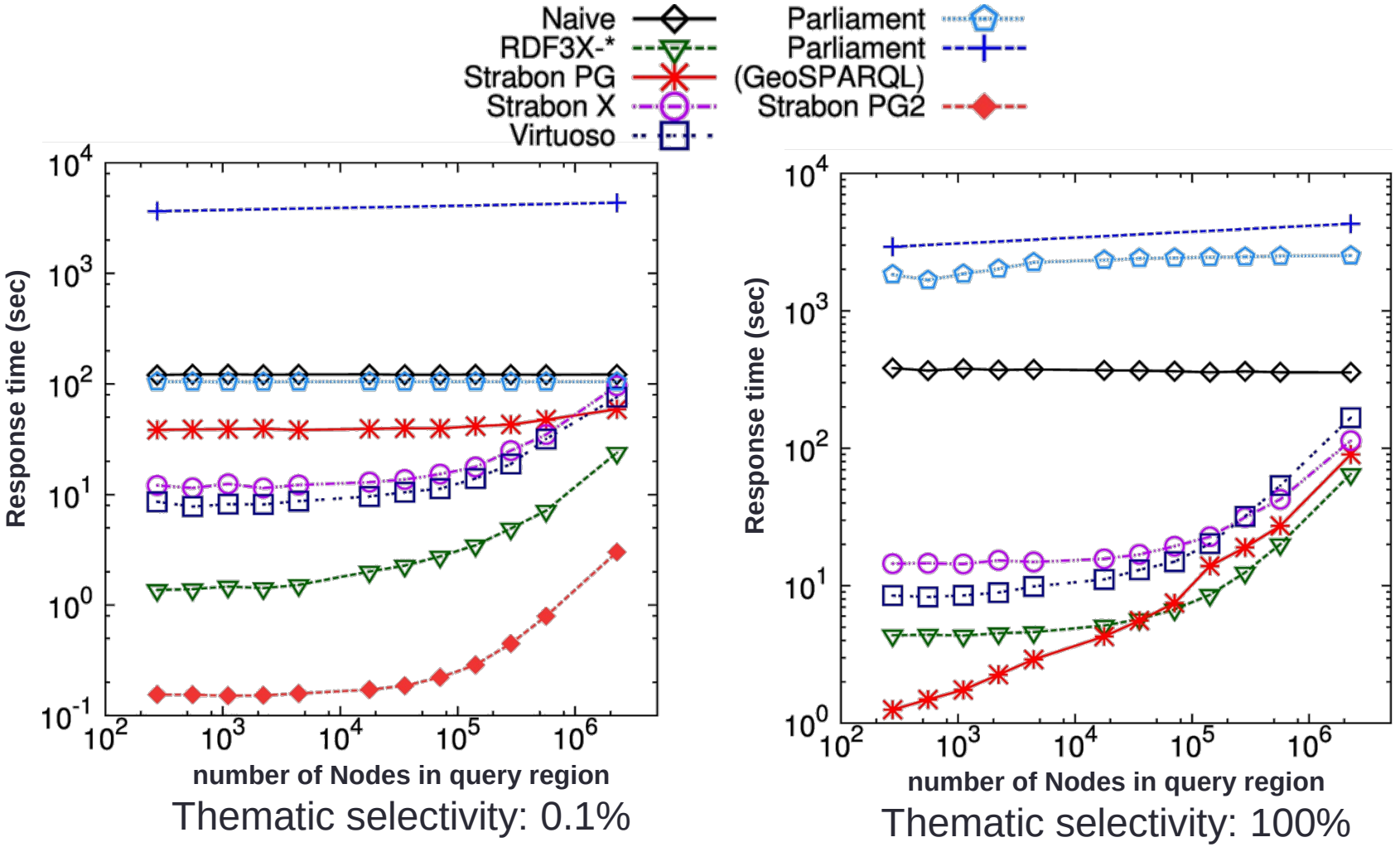
Strabon Architecture

[ISWC 2012
ESWC 2013]



Real-world Workload: 500 million triples – cold caches

[ISWC 2012]

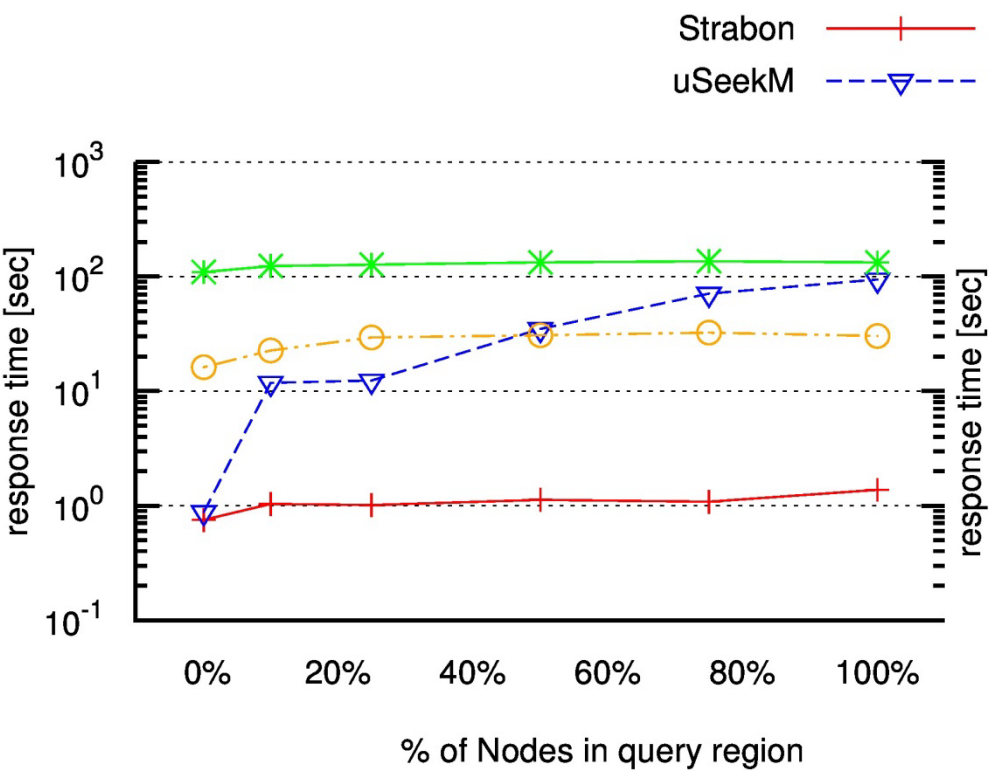


Thematic selectivity: 0.1%

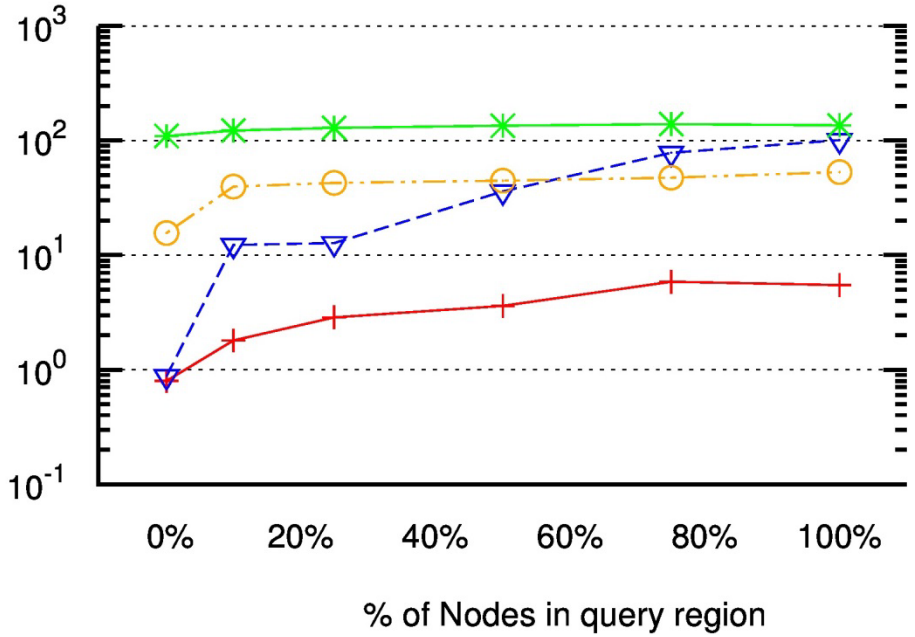
Thematic selectivity: 100%

Geographica
Synthetic Workload (Spatial Selections, cold caches)

[ISWC 2013]



Intersects
Thematic Selectivity: 0.2%



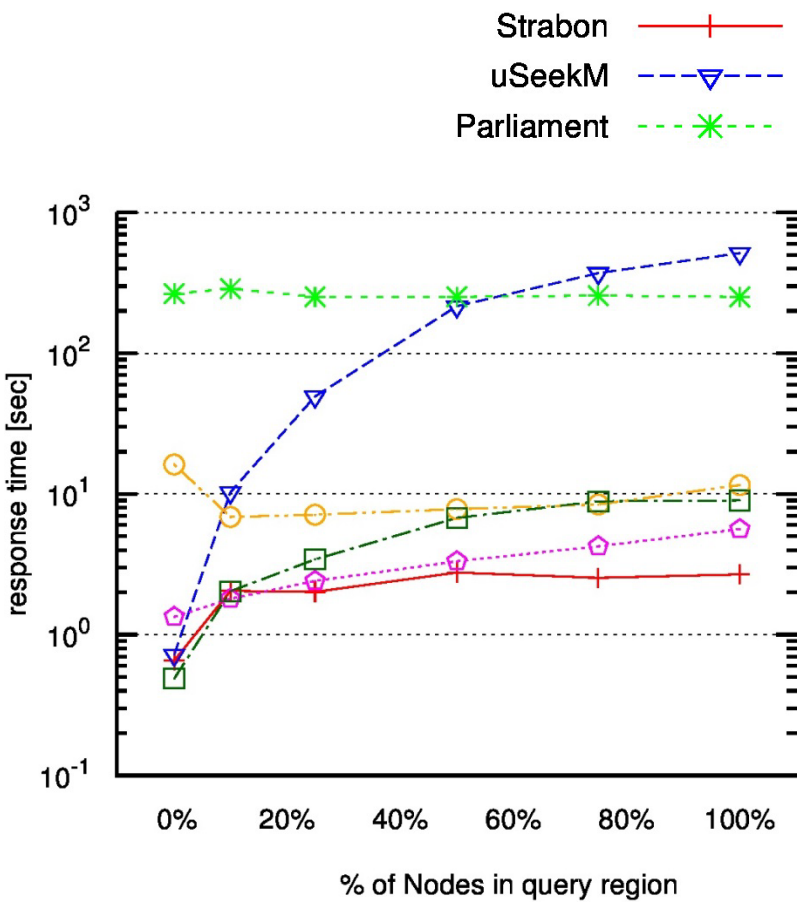
Intersects
Thematic Selectivity: 100%

http://geographica.di.uoa.gr

Results (points only)

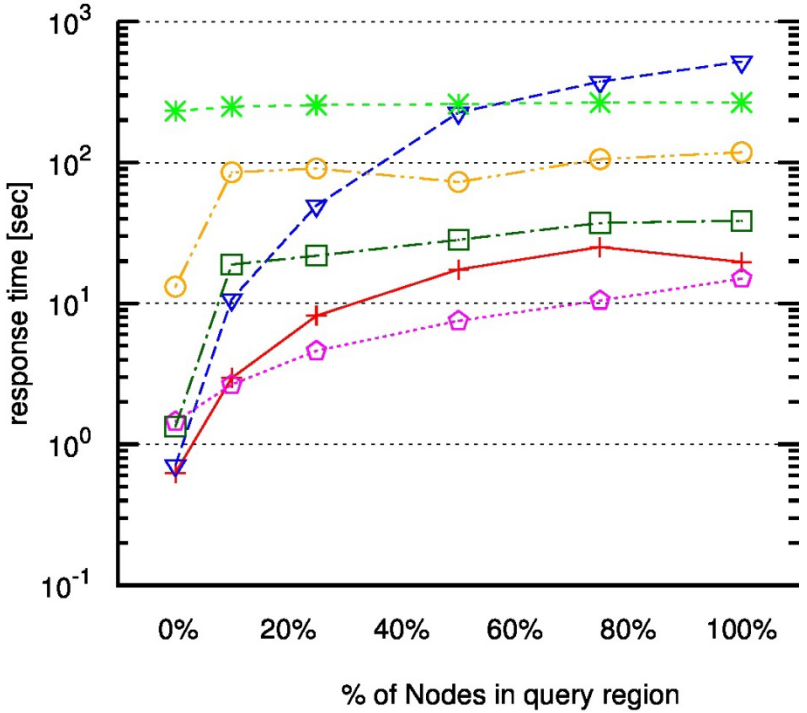
Synthetic Workload (Spatial Selections, cold caches)

[ISWC 2013]



Intersects

Thematic Selectivity: 0.1%

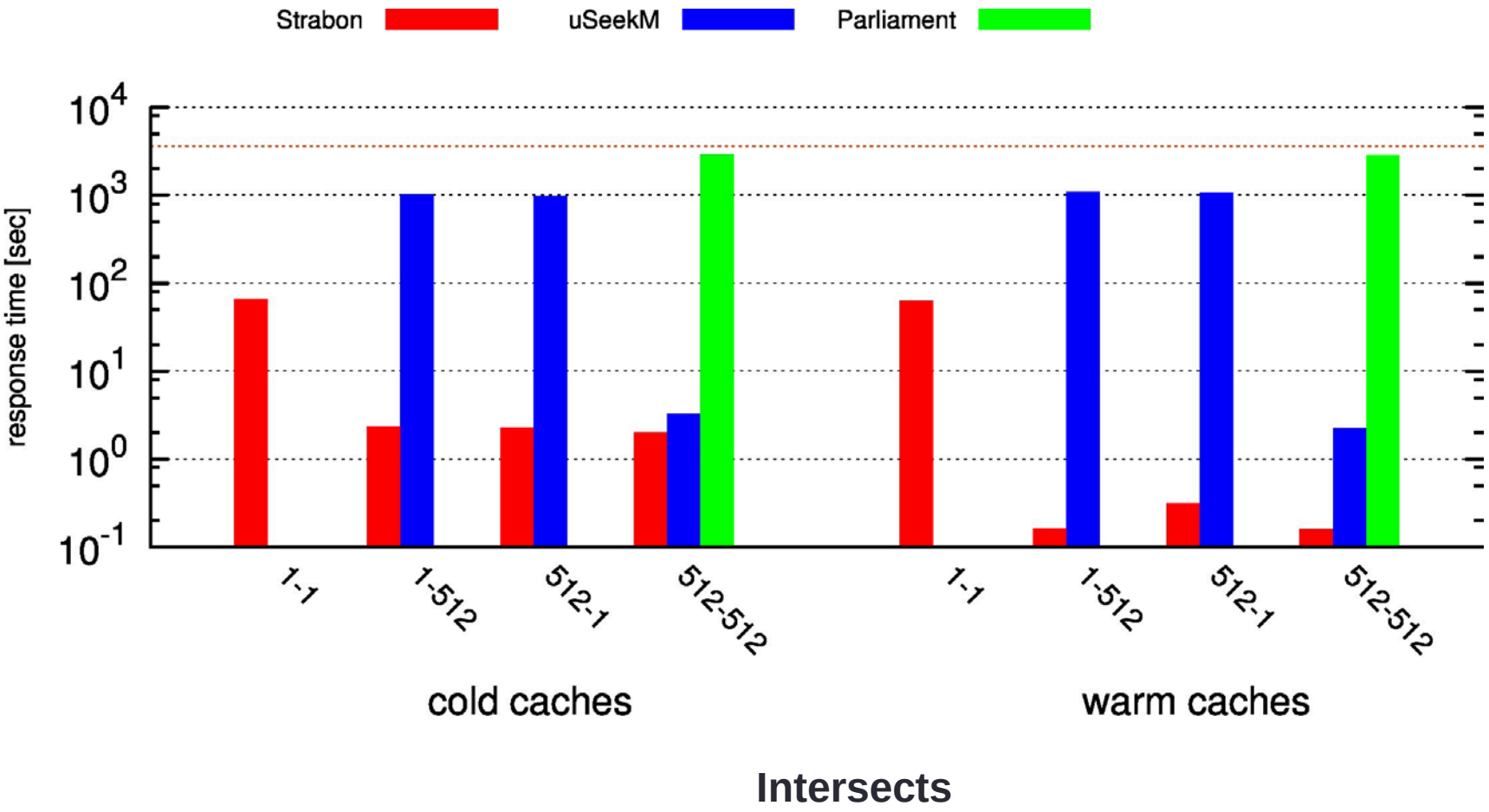


Intersects

Thematic Selectivity: 100%

Geographica
Synthetic Workload (Spatial Joins)

[ISWC 2013]



http://geographica.di.uoa.gr

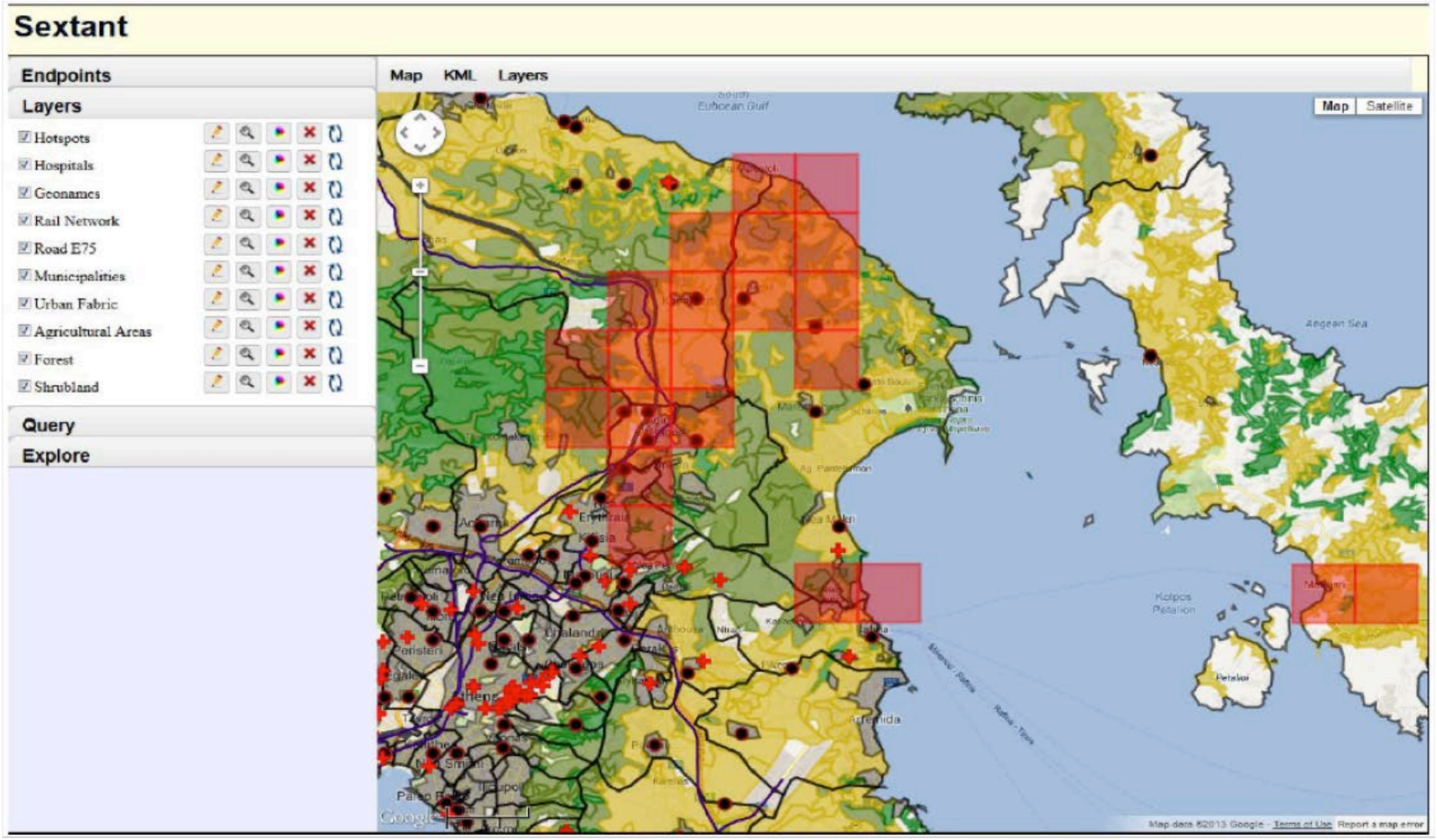
System	Language	Index	Geometries	CRS support	Geospatial Function Support
Strabon	stSPARQL/ GeoSPARQL*	R-tree-over-GiST	WKT / GML support	Yes	• OGC-SFA • Egenhofer • RCC-8
Parliament	GeoSPARQL*	R-Tree	WKT / GML support	Yes	•OGC-SFA •Egenhofer •RCC-8
Oracle	GeoSPARQL	R-Tree, Quadtree	WKT / GML support	Yes	•OGC-SFA •Egenhofer •RCC-8
Brodt et al. (RDF-3X)	SPARQL	R-Tree	WKT support	No	OGC-SFA
Perry	SPARQL-ST	R-Tree	GeoRSS GML	Yes	RCC-8
AllegroGraph	Extended SPARQL	Distribution sweeping technique	2D point geometries	Partial	•Buffer •Bounding Box •Distance
OWLIM	Extended SPARQL	Custom	2D point geometries	No	•Point-in-polygon •Buffer •Distance
Virtuoso	SPARQL	R-Tree	2D point geometries	Yes	SQL/MM (subset)
uSeekM	GeoSPARQL	R-tree-over GiST	WKT support	No	OGC-SFA

- The data model stRDF and the query language stSPARQL
- The system Strabon
- Visualizing time-evolving geometries using Sextant
- Real-Time Fire Monitoring application
- Conclusions

Visualizing time-evolving geometries using Sextant

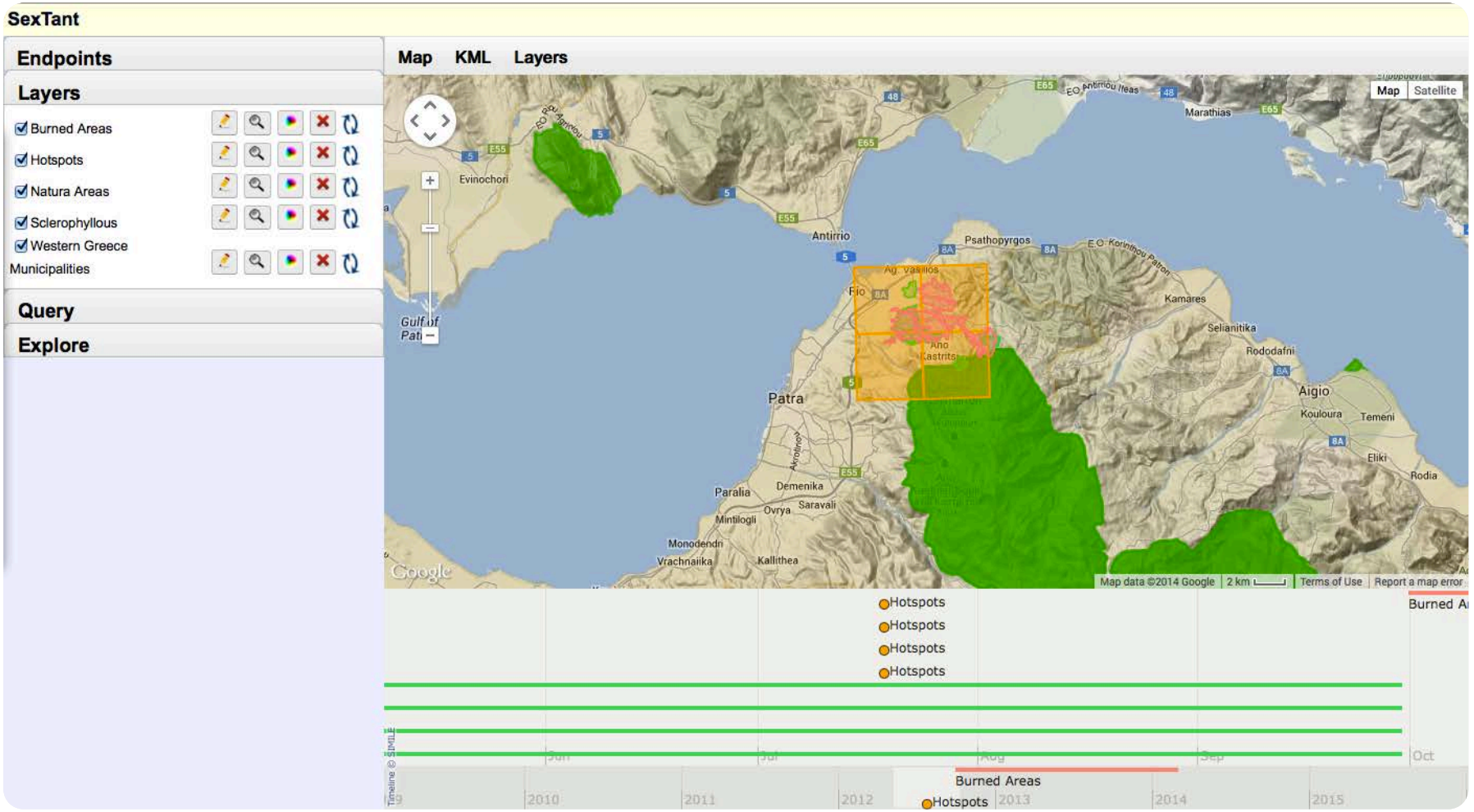
http://sextant.di.uoa.gr

Rapid Mapping application



http://bit.ly/sextant-rapid-mapping-attica

Evolution of Land Cover



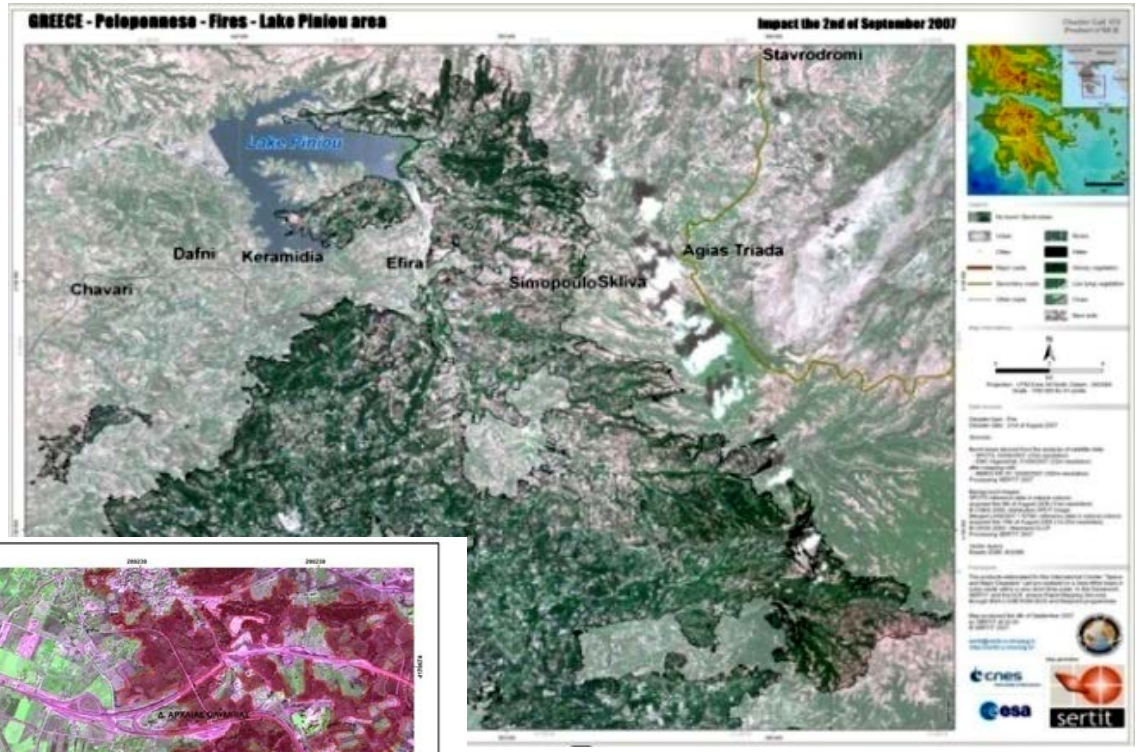
http://bit.ly/sextant-land-cover-evolution

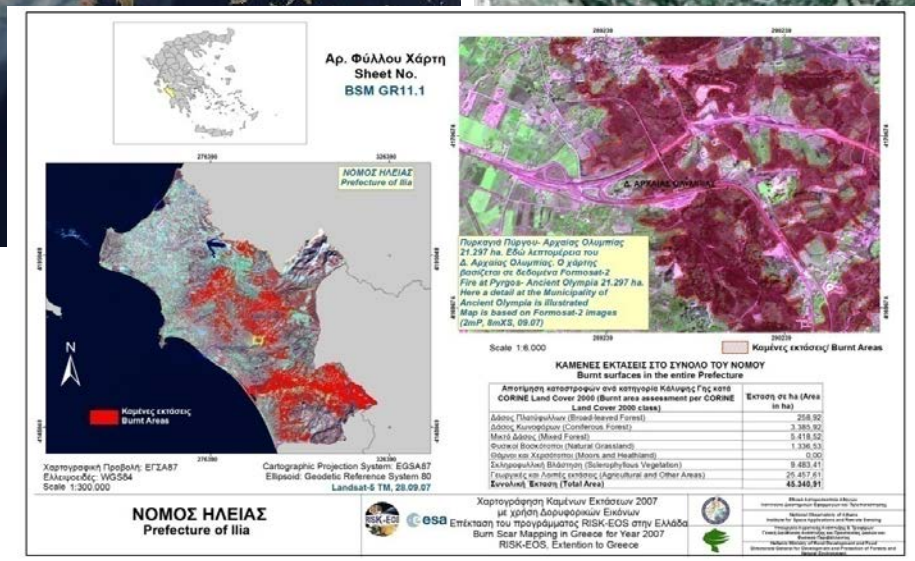
- The data model stRDF and the query language stSPARQL
- The system Strabon
- Visualizing time-evolving geometries using Sextant
- Real-Time Fire Monitoring application
- Conclusions

Real-Time Fire Monitoring

http://bit.ly/FiresInGreece

Wildfire Monitoring and Burnt Area Mapping (NOA)





EMERGENCY RESPONSE SERVICE GMES

INSTITUTE FOR SPACE APPLICATIONS AND REMOTE SENSING

High Level Data Modeling

- Need for representing
  - Standard product **metadata**
  - Standard product **semantic annotations**
  - **Geospatial information**
  - **Temporal information**

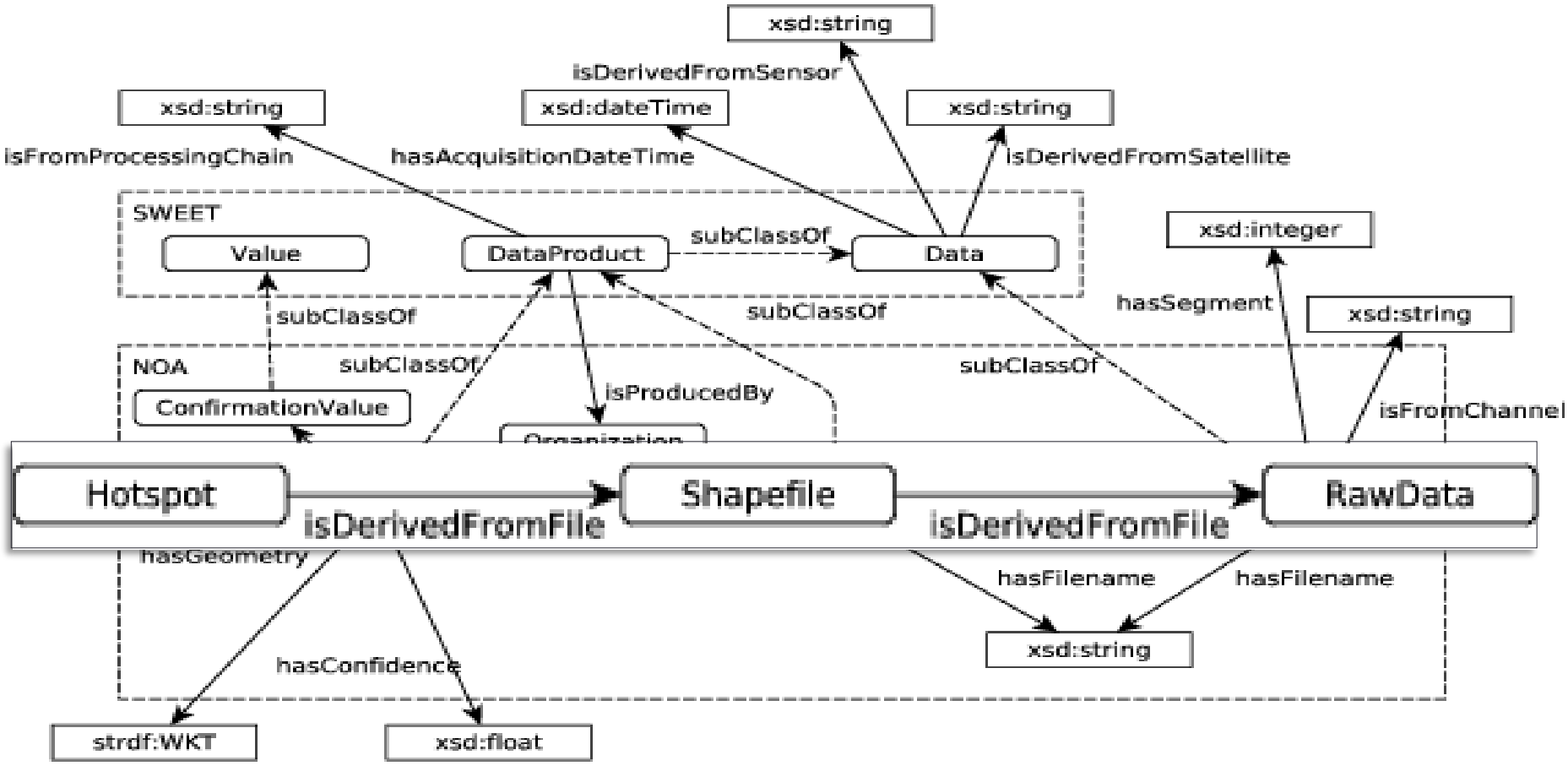
- Need to link to other data sources
  - **GIS data**
  - Other information on the **Web**

Fire Monitoring Application

- Improving the fire monitoring service using Semantic Web technologies
  - **Representing** fire related products using ontologies
  - **Enriching products** with linked geospatial data
  - **Improving accuracy** with respect to:
    - Underlying land cover/land use
    - Persistence in time

[ISWC 2012 Semantic Web Challenge 3'rd place winner]

NOA Ontology



Linked Geospatial Data

- Datasets that we developed and published as linked data:
  - Corine Land Use / Land Cover
  - Coastline of Greece
  - Greek Administrative Geography
- Portal: **http://www.linkedopendata.gr/**
- Datasets from Linked Open Data Cloud
  - LinkedGeoData
  - GeoNames

Linked Open Data

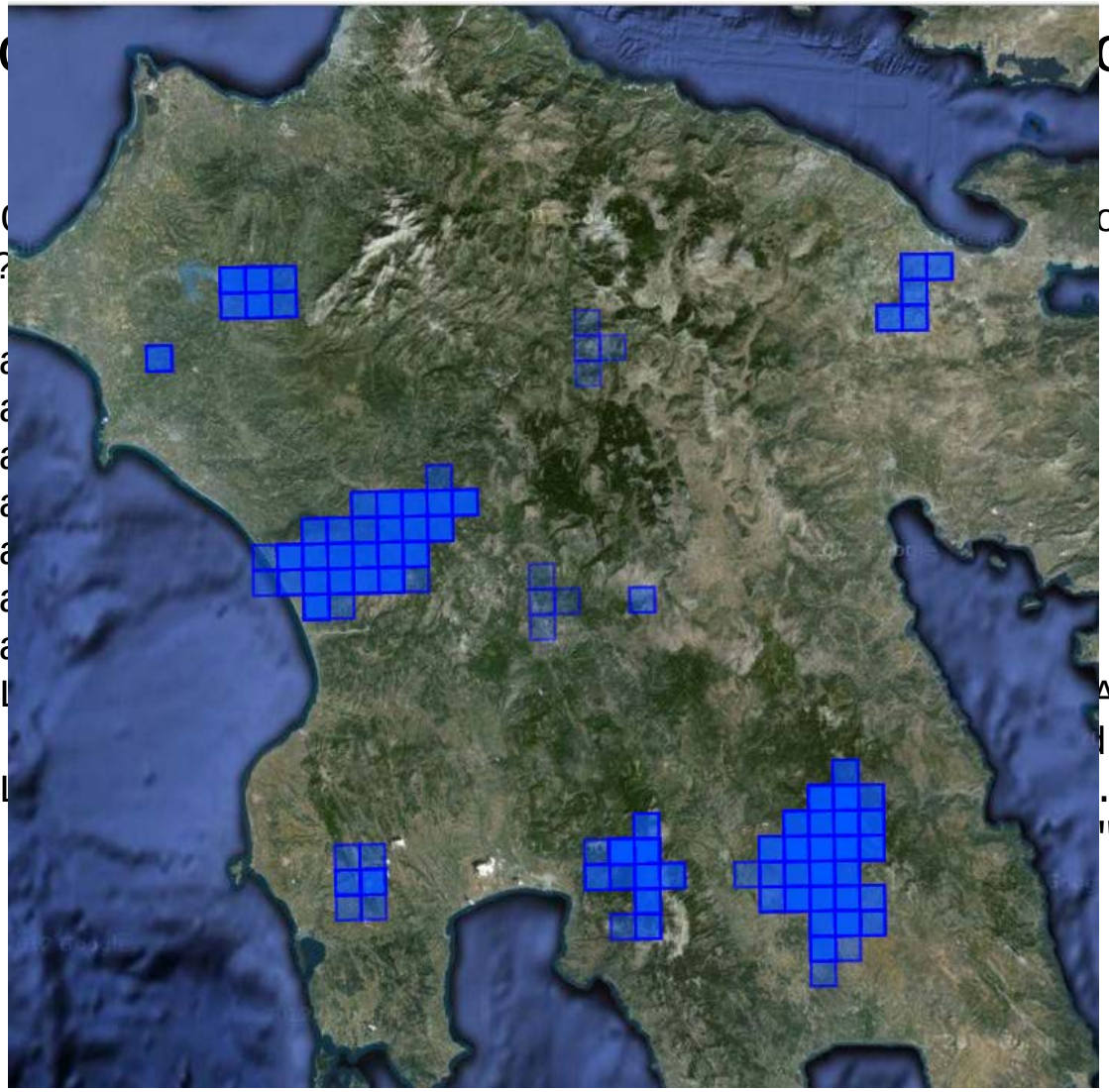
Improvements

Using ontologies and stRDF to model knowledge extracted from satellite images, metadata of satellite images and auxiliary geospatial data can improve tasks like:

- **Generated maps** combining diverse information sources
- **Increase hotspot accuracy** correlating them with auxiliary data

Discovering EO Data

Get all h on
24/08/



```
SELECT ?h ?h                                                    ovider
   ?hSensor ?
WHERE {   ?h
          noa
          noa
          noa
          noa
          noa
          noa
          noa
          FIL                                                   AcqTime &&
                                                                ateTime).
          FIL                                                   .77 38.36,
                                                                "
```

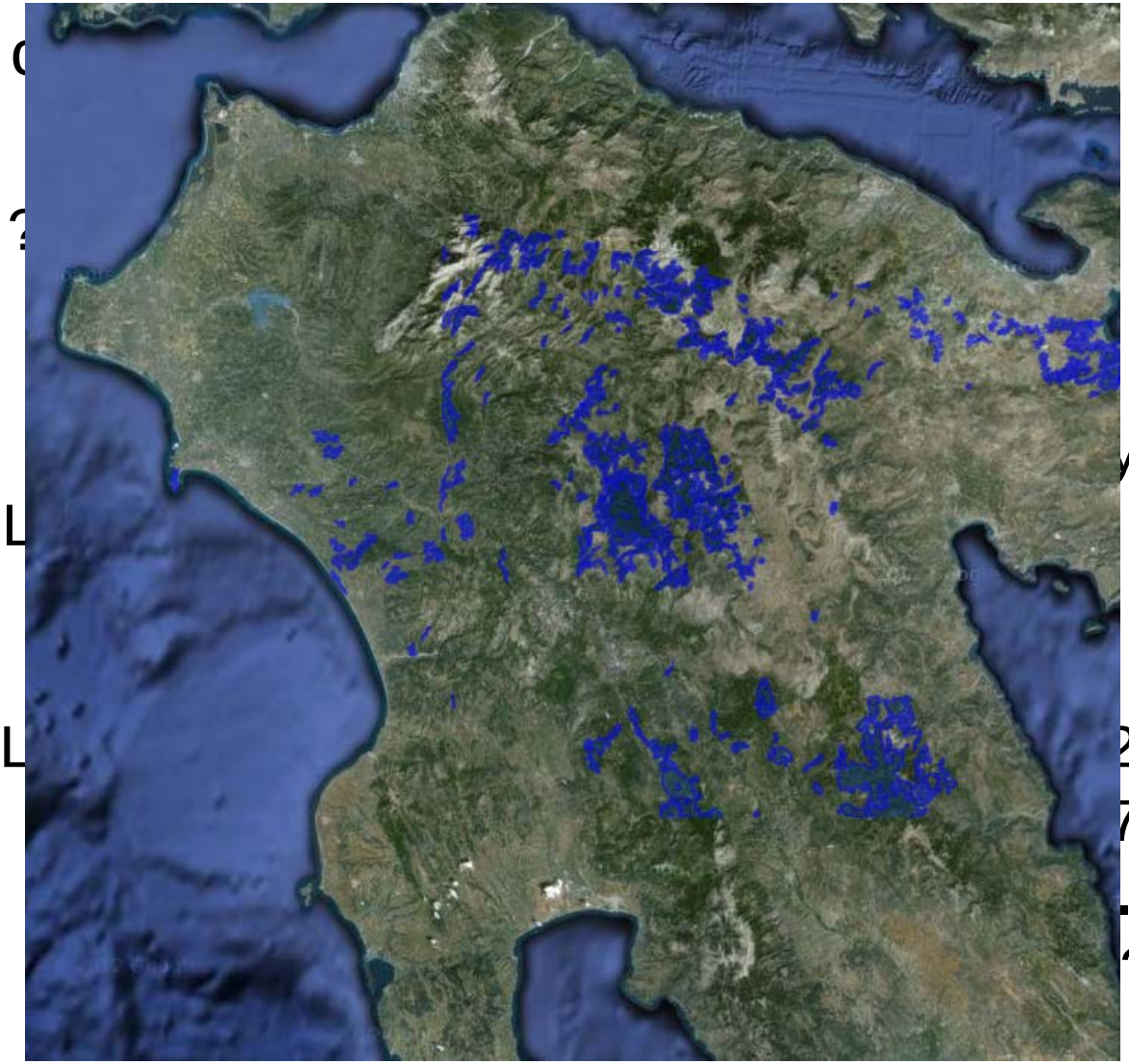
Retrieving a Map Layer (1/2)

Get all c

```
SELECT
WHERE{?

                                                    ype.
      FIL

                                                    .
      FIL                                           21.027
                                                    77 36.05,
                                                    .36))”
                                                    ?aGeo)).
}
```

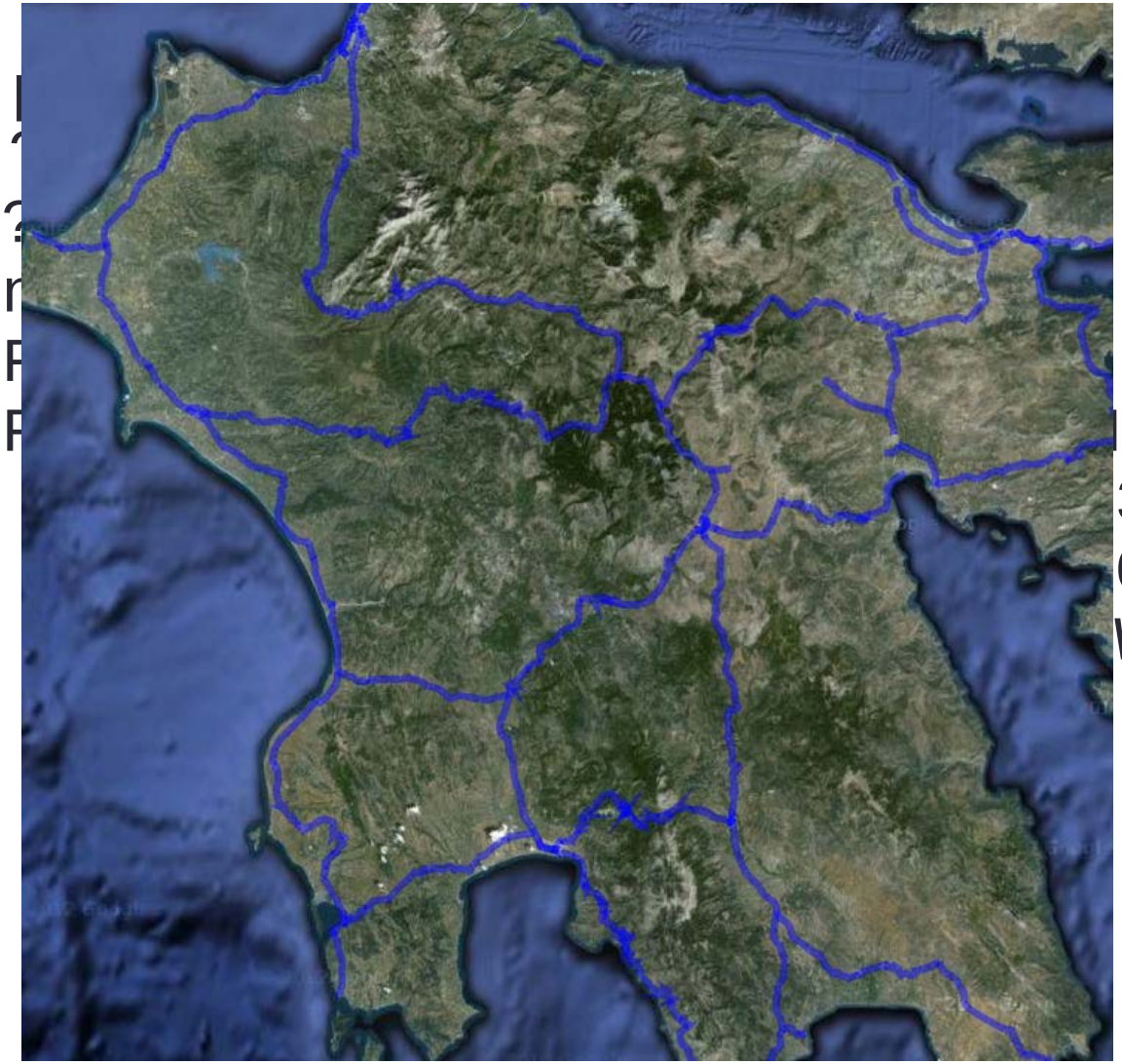


Retrieving a map layer (2/2)

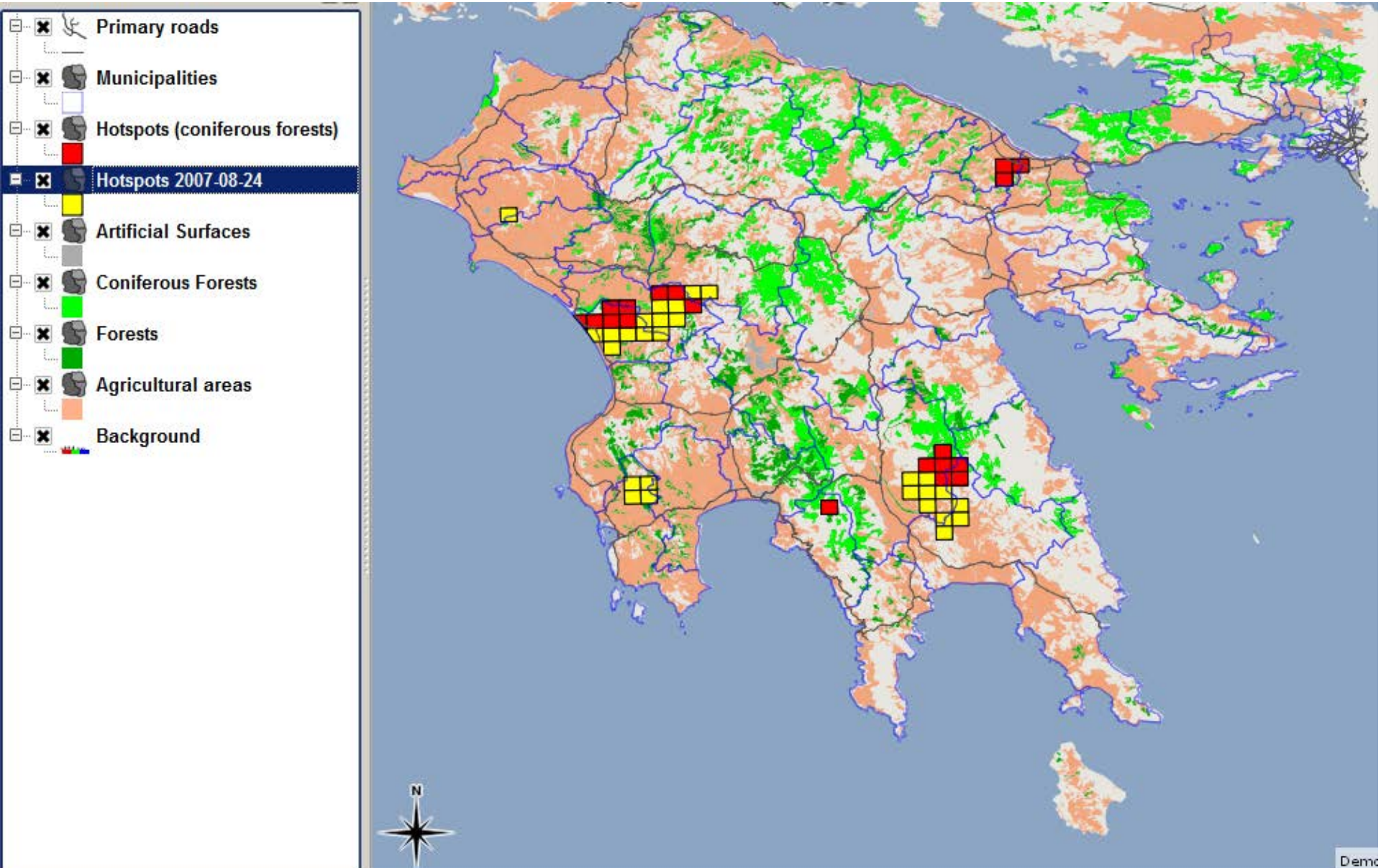
- Get all

```
SELECT
WHERE{ ?
r
F
F
                                                    .
                                                    ((
                                                    36,
                                                    05,
                                                    WKT,

}
```



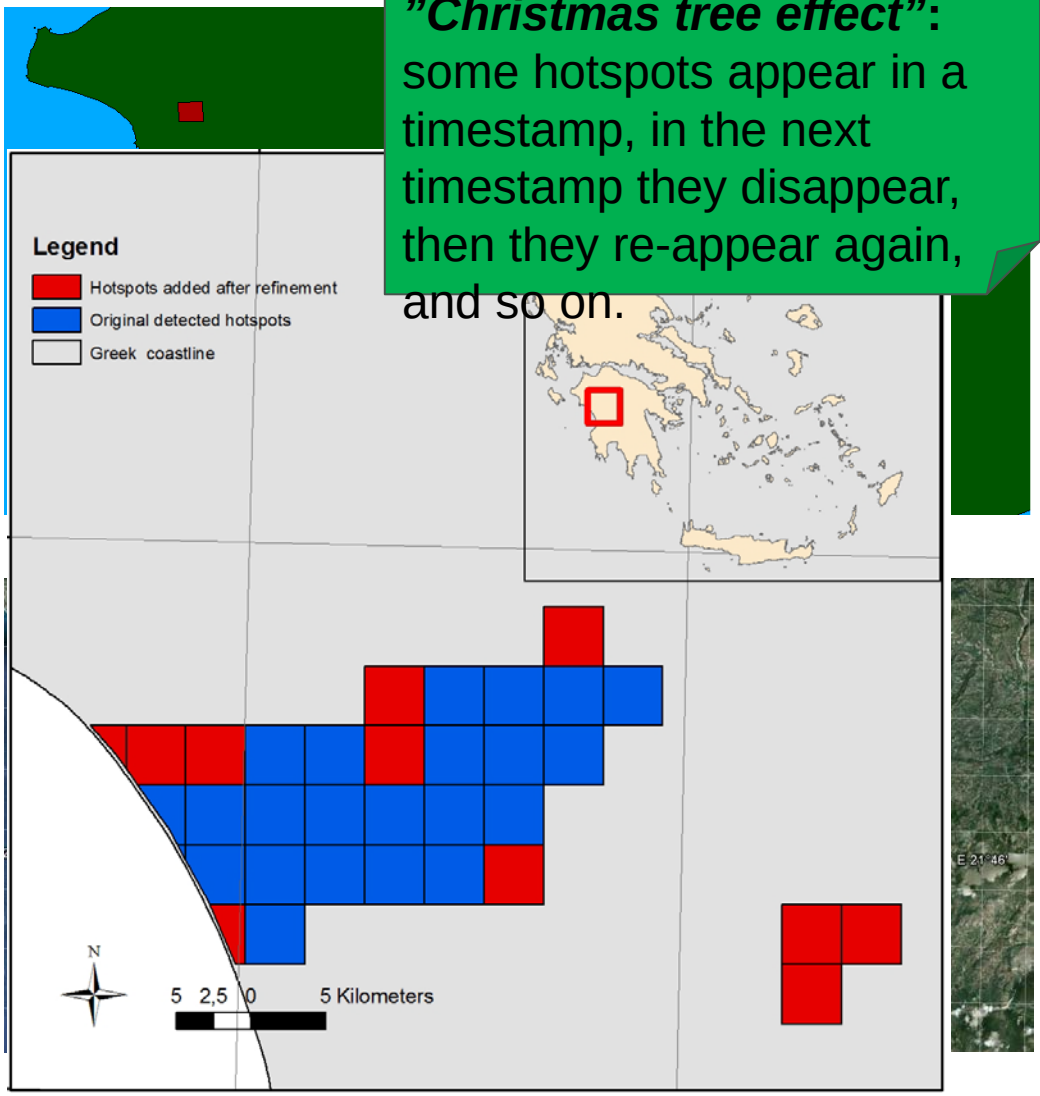
Final Map


Primary roads
Municipalities
Hotspots (coniferous forests)
Hotspots 2007-08-24
Artificial Surfaces
Coniferous Forests
Forests
Agricultural areas
Background
N

Semantic Enrichment for Hotspots

- **Enrich** hotspot products
  1. Connect each hotspot with a municipality that it is located
- **Improve accuracy** with respect to **underlying area**
  2. Eliminate false alarms in sea
  3. Eliminate false alarms in inconsistent land cover areas
  4. Keep land part of the polygon
- **Improve accuracy** with respect to **temporal persistence** of each hotspots
  5. Remove “Christmas tree” effects

***”Christmas tree effect”*:** some hotspots appear in a timestamp, in the next timestamp they disappear, then they re-appear again, and so on.



Conclusions

- The data model stRDF and the query language stSPARQL
- The system Strabon
  - The benchmark Geographica
- The tool Sextant for visualizing time-evolving geometric infromation
- General architecture for EO applications enriched with semantic web technologies

Vragen?

Kostis Kyzirakos
kostis@cwi.nl

Backup Slides

stSPARQL and GeoSPARQL

- GeoSPARQL is a recent effort by OGC to develop an extension of SPARQL for querying geospatial data expressed in RDF.
- stSPARQL and GeoSPARQL have been developed independently.
- Functionalities **similar to stSPARQL**:
  - Geometries are represented using literals similarly to stSPARQL.
  - The same families of functions are offered for querying geometries.
- Functionalities **beyond stSPARQL**:
  - Topological relations can now be asserted as well so that reasoning and querying on them is possible.
- However, GeoSPARQL does not have well-defined semantics, does not discuss modeling and querying of temporal information, nor it offers spatial aggregates.
- Strabon supports both stSPARQL and GeoSPARQL.