Building Virtual Earth Observatories Using Semantic Web Technologies

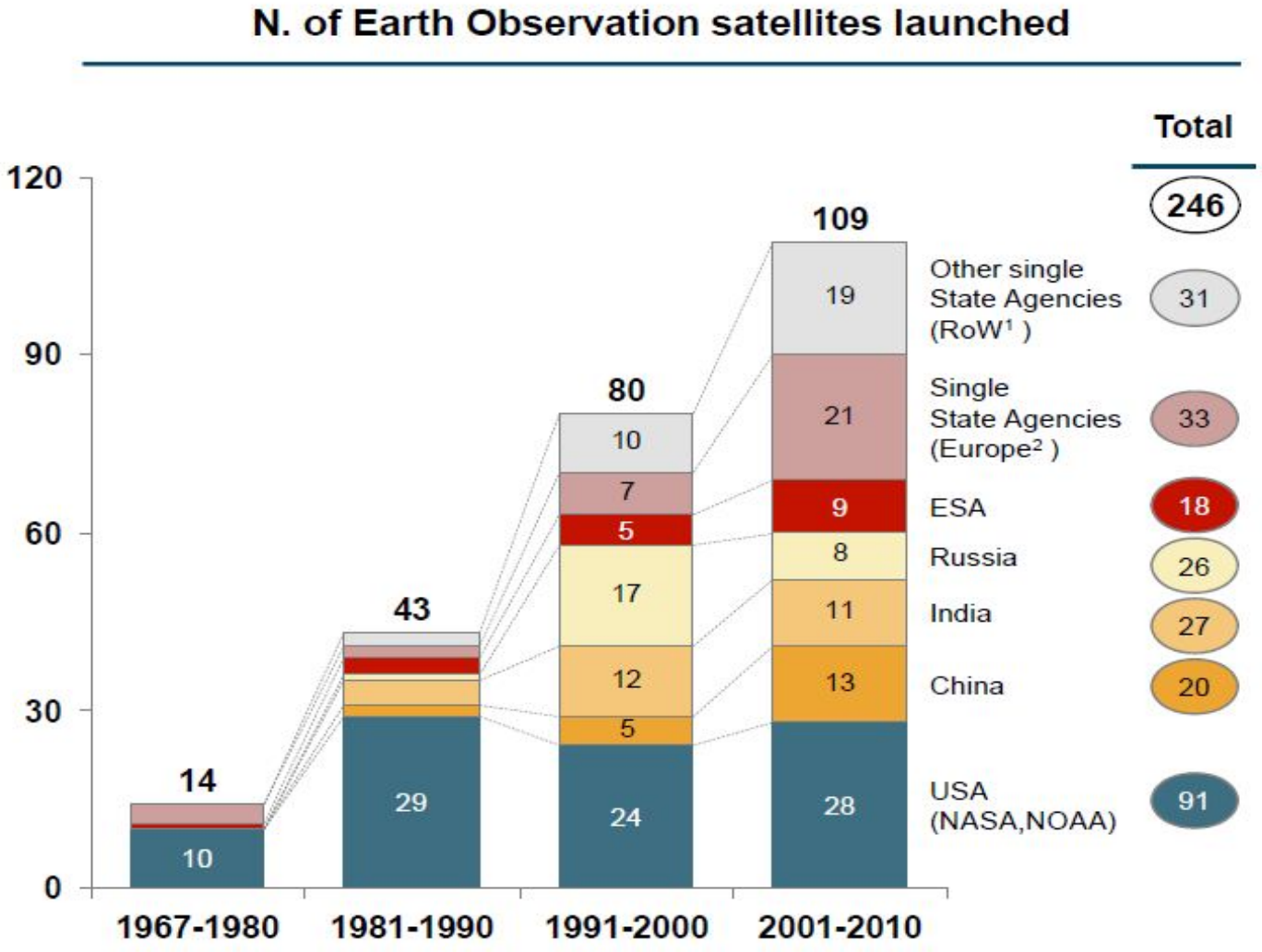
Manolis Koubarakis
Dept. of Informatics and Telecommunications
National and Kapodistrian University of Athens

Inria visit May 2012 2012

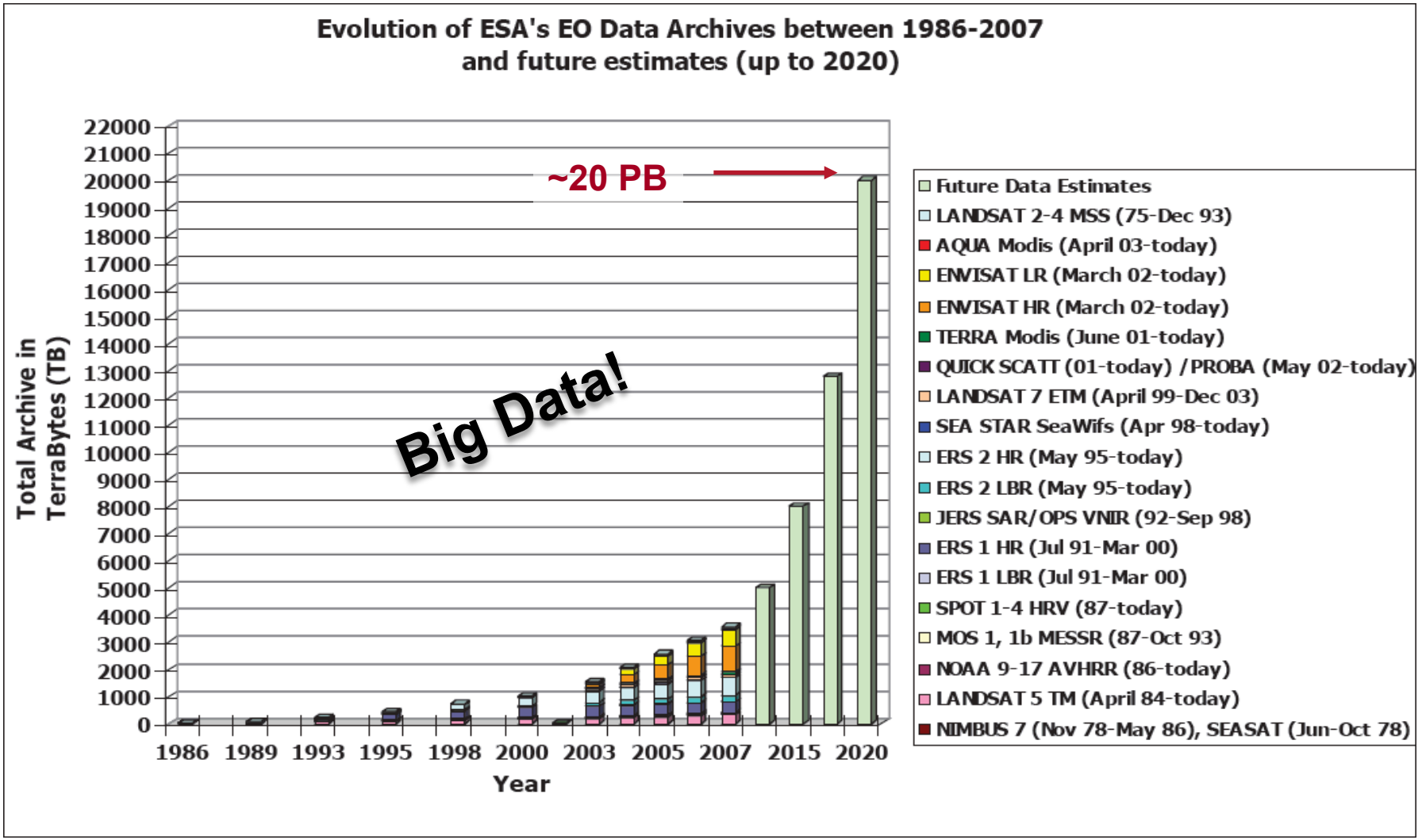
Outline

- Motivation for project TELEIOS
- State of the art in Earth Observation data centers
- Developing Virtual Earth Observatories using Semantic Web technologies
- Technical highlights of the work of my group in this area.

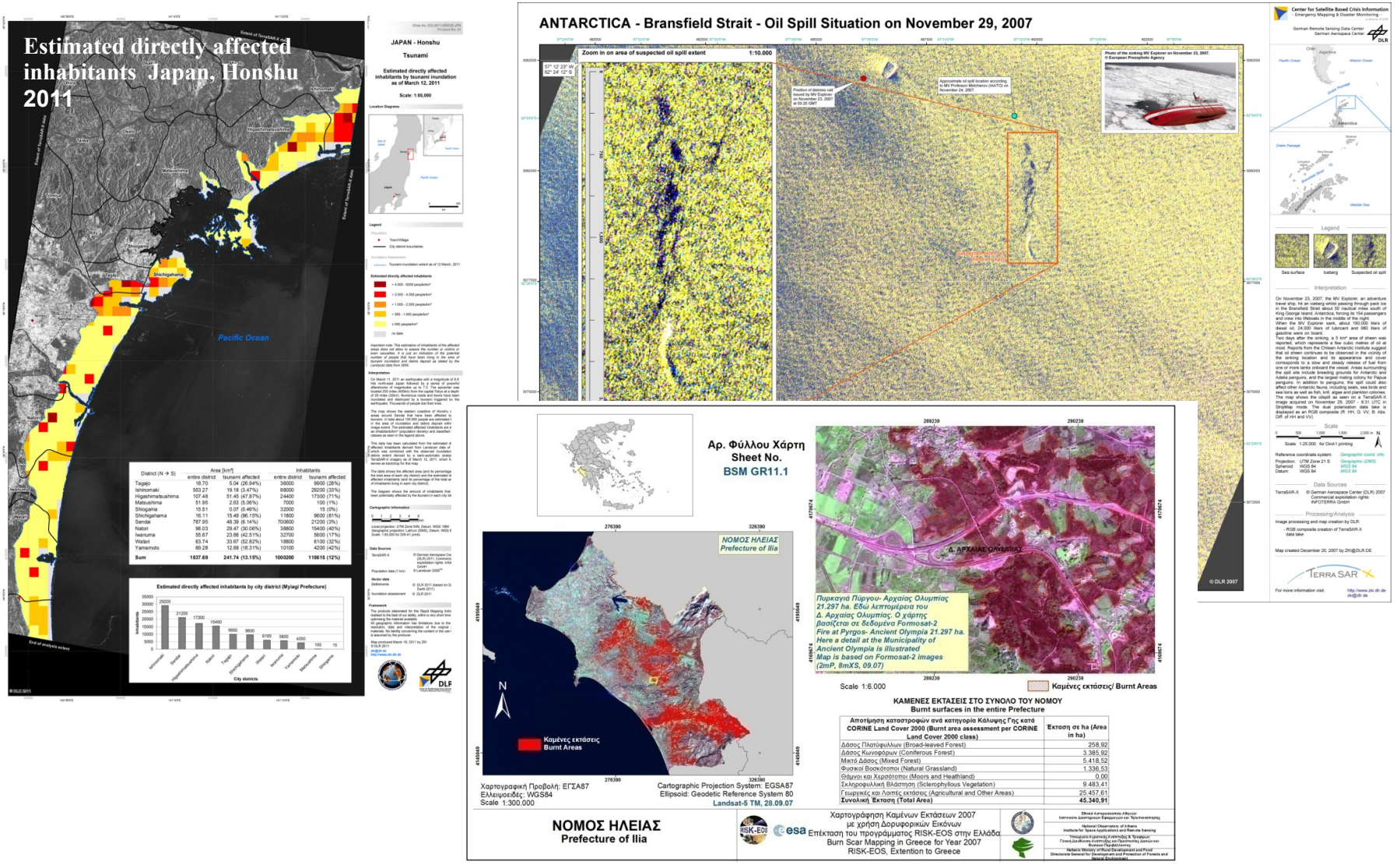
Motivation


N. of Earth Observation satellites launched
Total
246
120
90
60
30
0
109
80
43
14
19
21
10
7
5
17
12
5
29
24
10
28
9
8
11
13
Other single State Agencies (RoW¹)
31
Single State Agencies (Europe²)
33
ESA
18
Russia
26
India
27
China
20
USA (NASA,NOAA)
91
1967-1980
1981-1990
1991-2000
2001-2010

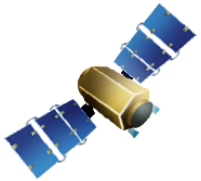
Motivation (cont'd)



Motivation (cont'd)



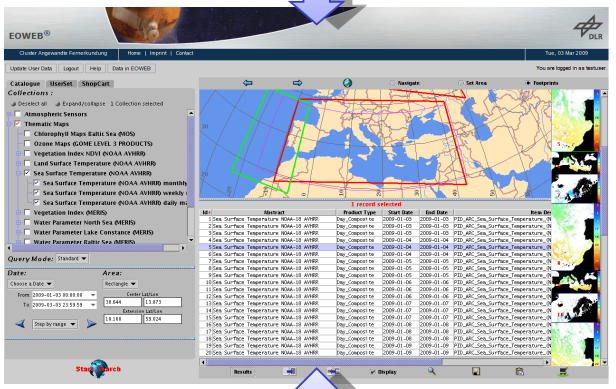
State of the Art in EO Data Centers



Raw Data

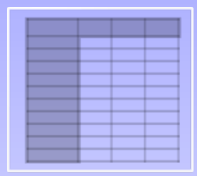
Users

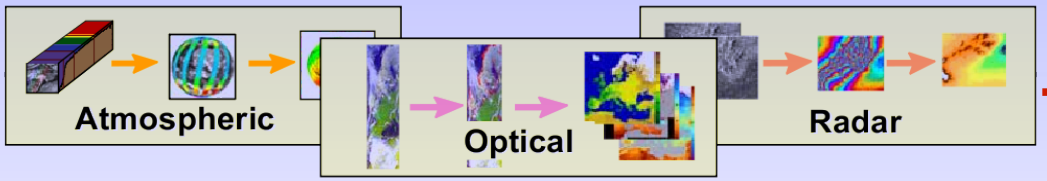
EOWEB



EO data center

Catalogue





Processing Chains

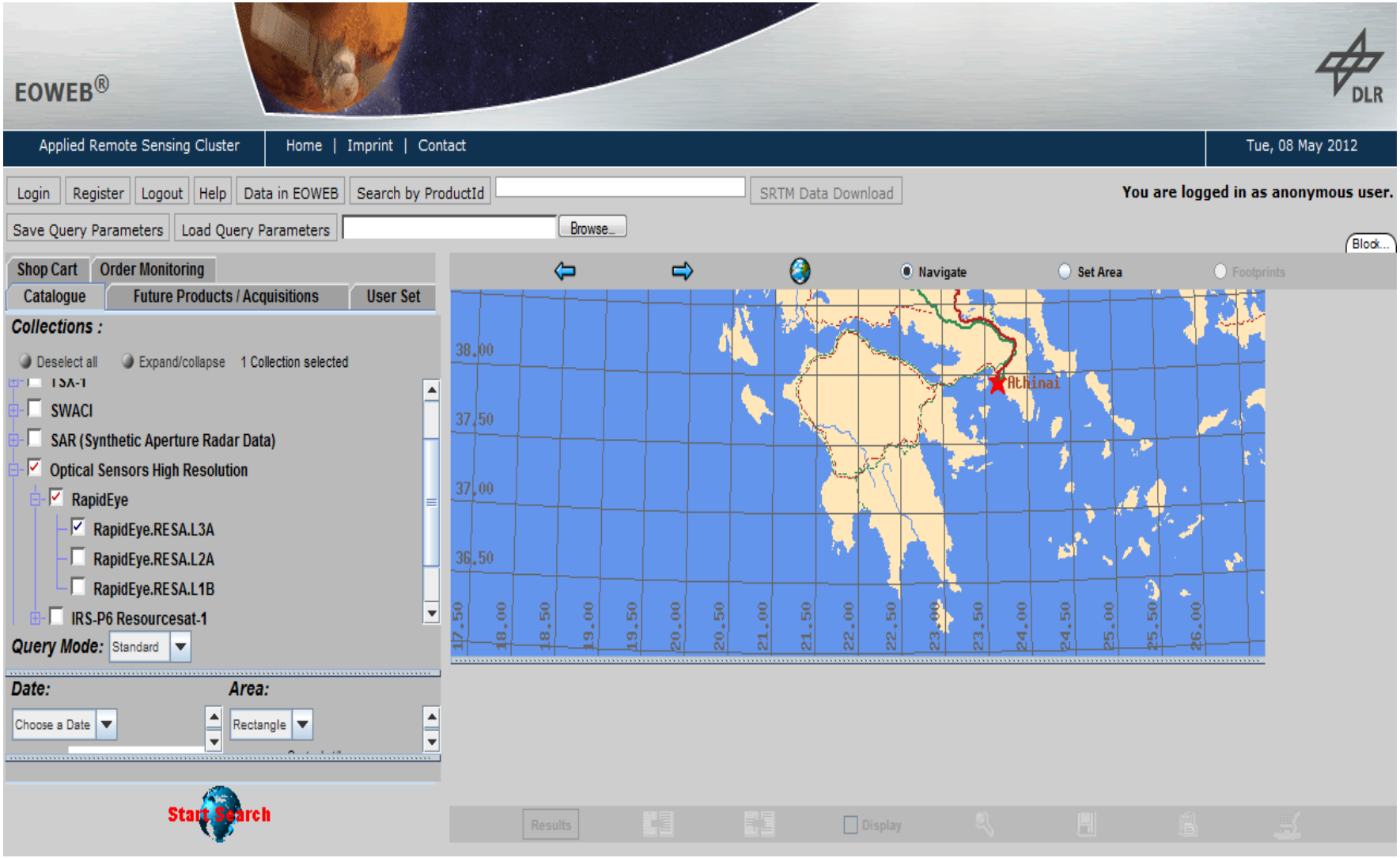
Archive

Example

- Can I pose the following query using EOWEB?

  Find images taken by the MSG2 satellite on August 25, 2007 which contain fire hotspots in areas which have been classified as forests according to Corine Land Cover, and are located within 2km from an archaeological site in the Peloponnese.

Example (cont'd)



Example (cont'd)

- Well, only partially.

  Find images taken by the MSG2 satellite on August 25, 2007 which contain fire hotspots in areas which have been classified as forests according to Corine Land Cover, **and are located within** 2km from an archaeological site in the **Peloponnese.**

Example (cont'd)

- But why?

- All this information is available in the **satellite images** and other **auxiliary data sources** of EO data centers or **on the Web.**

- However, EO data centers today do not allow:
  - **the mining of satellite image content** and
  - **its integration with other relevant data sources** so the previous query can be answered.

The TELEIOS Earth Observatory: Concept View

Scientists

Emergency Response Managers

Web Portals

Rapid Mapping

Scientific Database and Semantic Web Technologies

Linked Geospatial Data

Semantic Annotations

Ontologies

Knowledge Discovery and Data Mining

KNOWL-EDGE

GIS Data

Derived Products

Meta

Features

Raw Data

Ingestion

Proce

Catalo

Arch

Content Extraction

DATA

Data Modeling Requirements in EO Data Centers

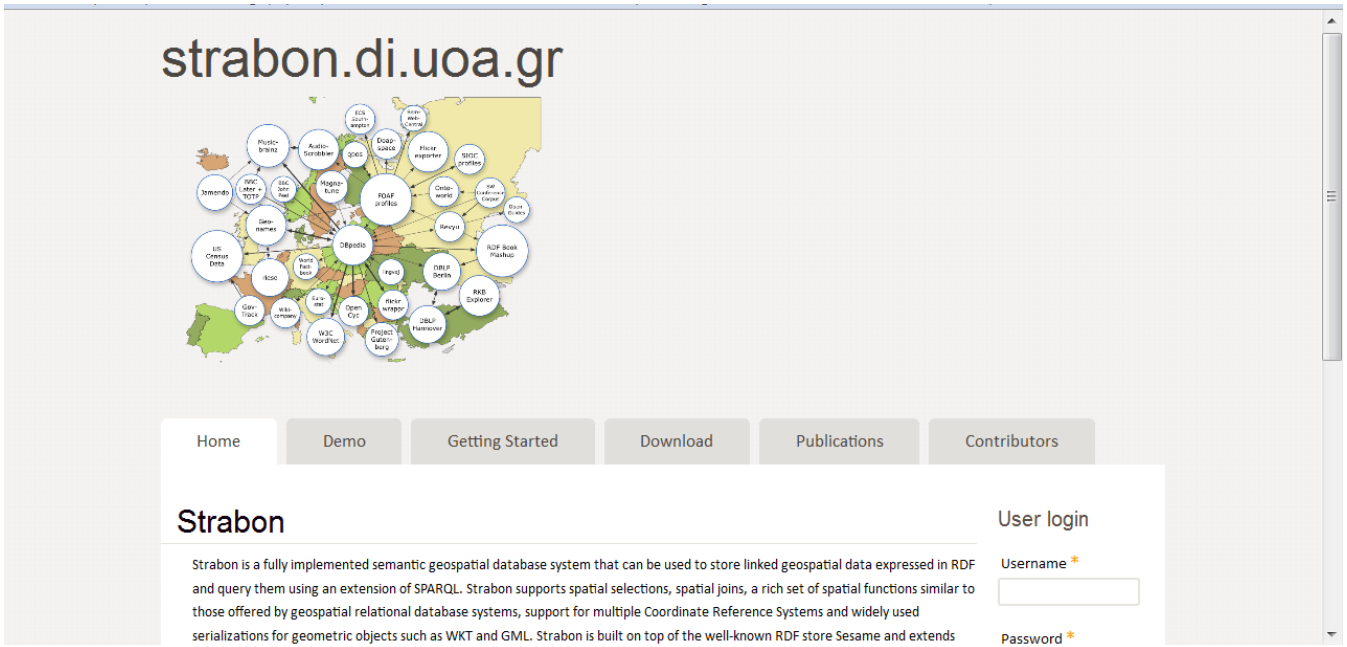
- Need for modeling
  - Standard product **metadata**
  - Standard product **content**
  - **Geospatial information**
  - **Temporal information**

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

Semantics-Based Representation and Querying of EO Data

- The data model **stRDF** and the query language **stSPARQL**

- The system **Strabon**



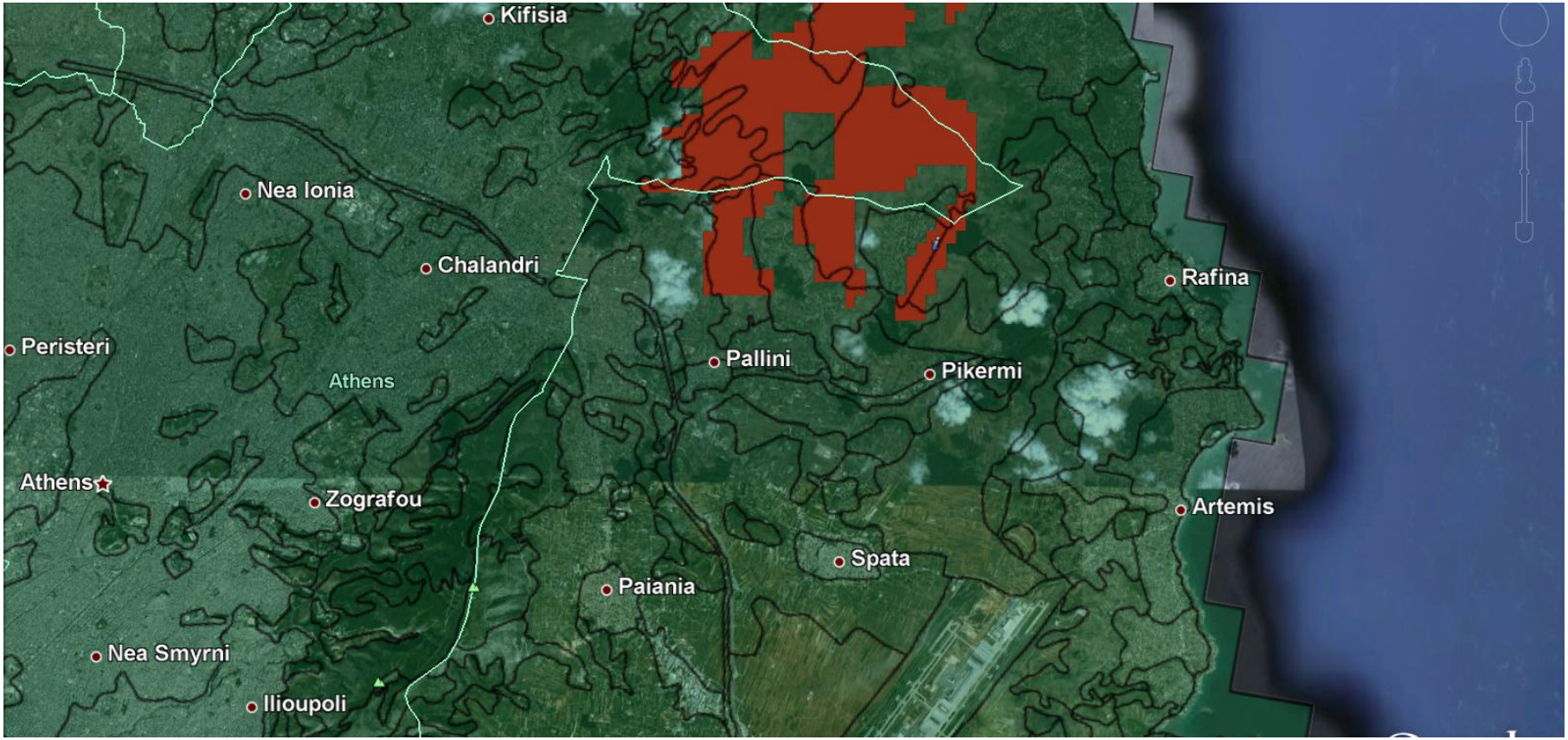
The Data Model stRDF

- **stRDF** stands for **spatiotemporal RDF.**

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

- 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`)
  - **Valid time of triples** (ignored in this talk)

stRDF: An example



stRDF: An example (WKT)

```
ex:BurntArea1 rdf:type noa:BurntArea.

ex:BurntArea1 noa:hasID "1"^^xsd:decimal.

ex:BurntArea1 noa:hasArea "23.7636"^^xsd:double.

ex:BurntArea1 strdf:hasGeometry "POLYGON(( 38.16 23.7, 38.18
  23.7,    38.18 23.8, 38.16 23.8, 38.16 23.7)); <http://
  spatialreference.org/ref/epsg/4121/>"^^strdf:WKT .
```

Spatial Literal (OpenGIS Simple Features)

Spatial Data Type Well-Known Text

stRDF: An example (GML)

```
ex:BurntArea1 rdf:type noa:BurntArea.
ex:BurntArea1 noa:hasID "1"^^xsd:decimal.
ex:BurntArea1 noa:hasArea "23.7636"^^xsd:double.

ex:BurntArea1 strdf:hasGeometry
       "<gml:Polygon
          srsName='http://www.opengis.net/def/crs/EPSG/0/4121'>
         <gml:outerBoundaryIs>
          <gml:LinearRing>
           <gml:coordinates>38.16,23.70 38.18,23.70 38.18,
                            23.80 38.16,23.80,38.16 23.70
           </gml:coordinates>
          </gml:LinearRing>
         </gml:outerBoundaryIs>
        </gml:Polygon>"^^strdf:GML .
```

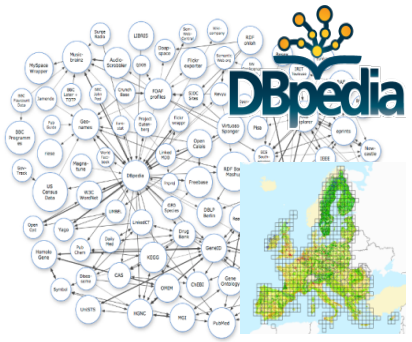
Spatial Literal (GML Simple Features Profile)

Spatial Data Type GML

stSPARQL: An example

- Find all burnt forests close to a city

```
SELECT ?BA ?BAGEO
WHERE {
        ?R rdf:type noa:Region .
        ?R strdf:hasGeometry ?RGEO .
        ?R noa:hasCorineLandCoverUse ?F .
        ?F rdfs:subClassOf clc:Forests .
        ?CITY rdf:type dbpedia:City .
        ?CITY strdf:hasGeometry ?CGEO .
        ?BA rdf:type noa:BurntArea .
        ?BA strdf:hasGeometry ?BAGEO .
        FILTER(strdf:anyInteract(?RGEO,?BAGEO) &&
               strdf:distance(?BAGEO,?CGEO) < 0.02) }
```



Spatial Functions

stSPARQL: More details

- We start from **SPARQL 1.1.**
- We add the property `strdf:hasGeometry` for **connecting RDF resources to spatial literals.**
- We add a **SPARQL extension function** for each function defined in the OGC standard **OpenGIS Simple Feature Access – Part 2: SQL option (ISO 19125)** for adding geospatial data to relational DBMSs and SQL.

stSPARQL (cont'd)

- **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`)

stSPARQL (cont'd)

- **Spatial terms**

  Constants (e.g., `"POLYGON((38.16 23.7, ...)) " ^^strdf:WKT`)

  Variables (e.g., `?GEO`)

  Results of set operations (e.g., `strdf:intersection`, `strdf:union`)

  Results of geometric operations (e.g., `strdf:boundary`, `strdf:buffer`)

- **SELECT clause**

  Construction of new geometries (e.g., `strdf:buffer(?geo, 0.1)`)

  Spatial aggregate functions (e.g., `strdf:extent(?geo)`)

  Metric functions (e.g., `strdf:area(?geo)`)

- **FILTER clause**

  Functions for testing topological spatial relationships between spatial terms (e.g., `strdf:contains(?G1, strdf:union(?G2, ?G3))`)

  Numeric expressions involving spatial metric functions (e.g.,`strdf:area(?G1)<=2*strdf:area(?G2)+1`)

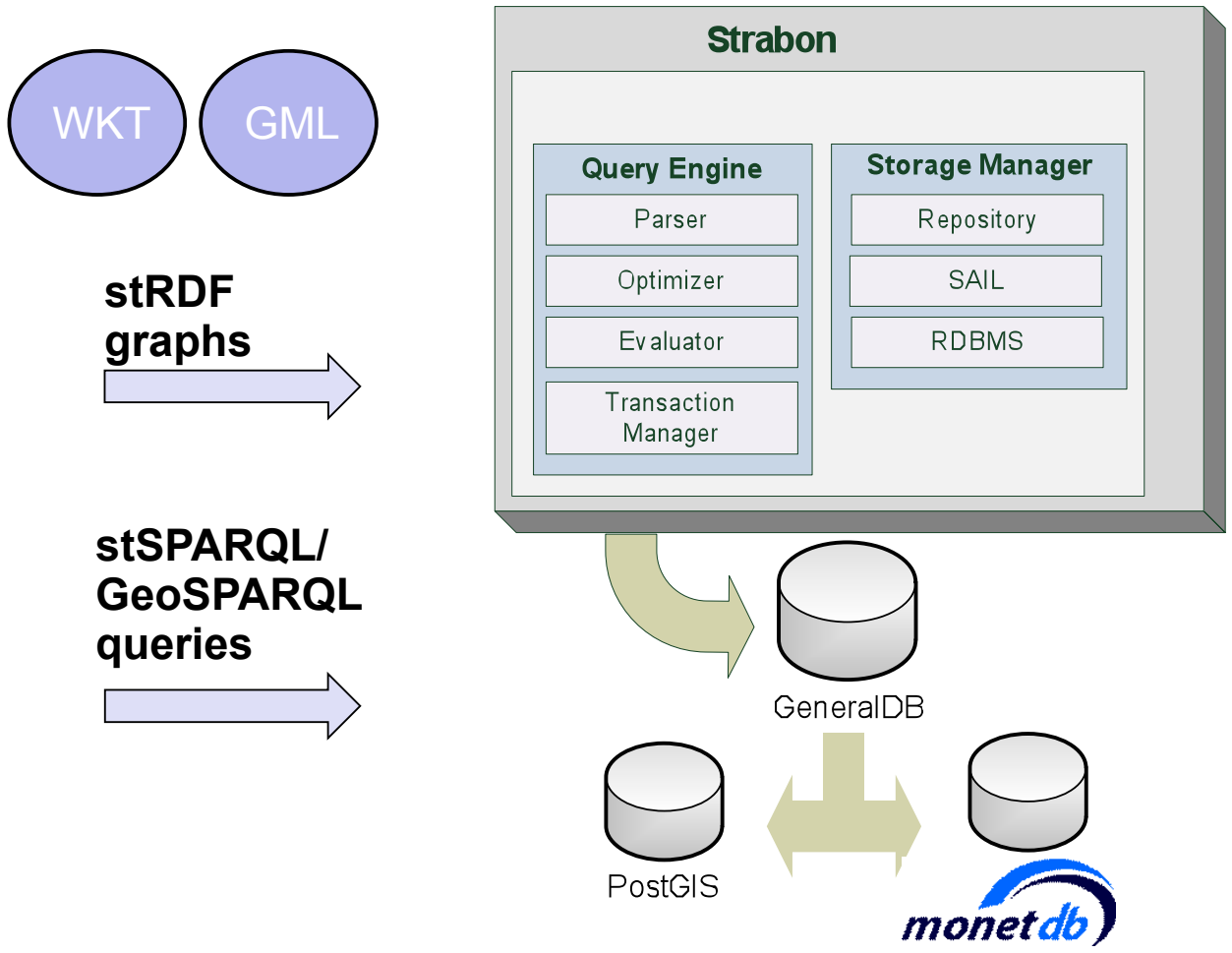
- **HAVING clause**

  Spatial aggregate functions and spatial metric functions or functions testing for topological relationships between spatial terms (e.g., `strdf:area(strdf:union(?geo))>1`)

stSPARQL (cont'd)

- stSPARQL geospatial query functionality is essentially a **syntactic alternative to a subset of the recent OGC standard GeoSPARQL (Strabon supports both syntaxes)**
- We add appropriate geospatial extensions to **SPARQL 1.1 Update language**

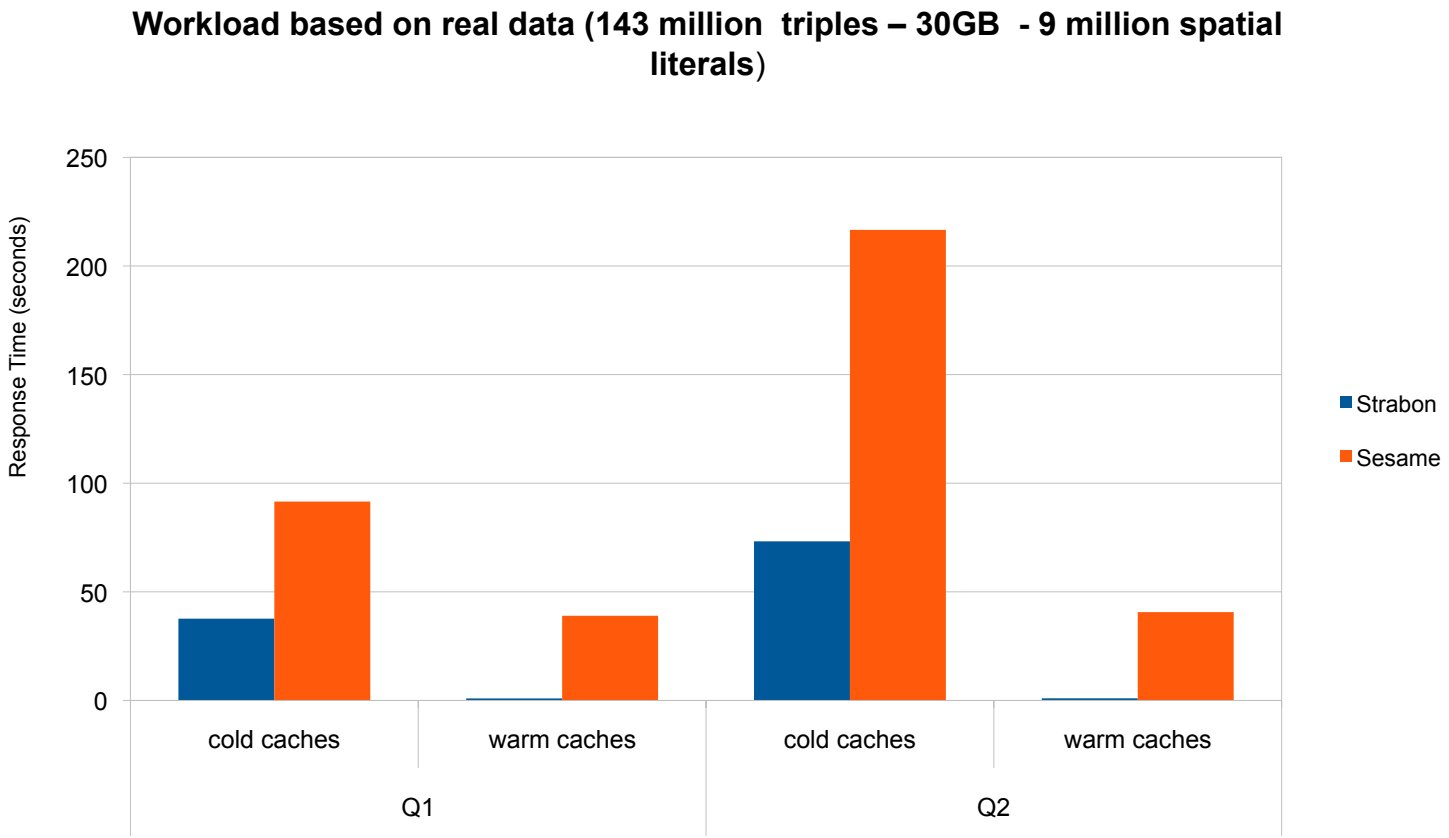
Strabon: A Scalable Geospatial RDF Store



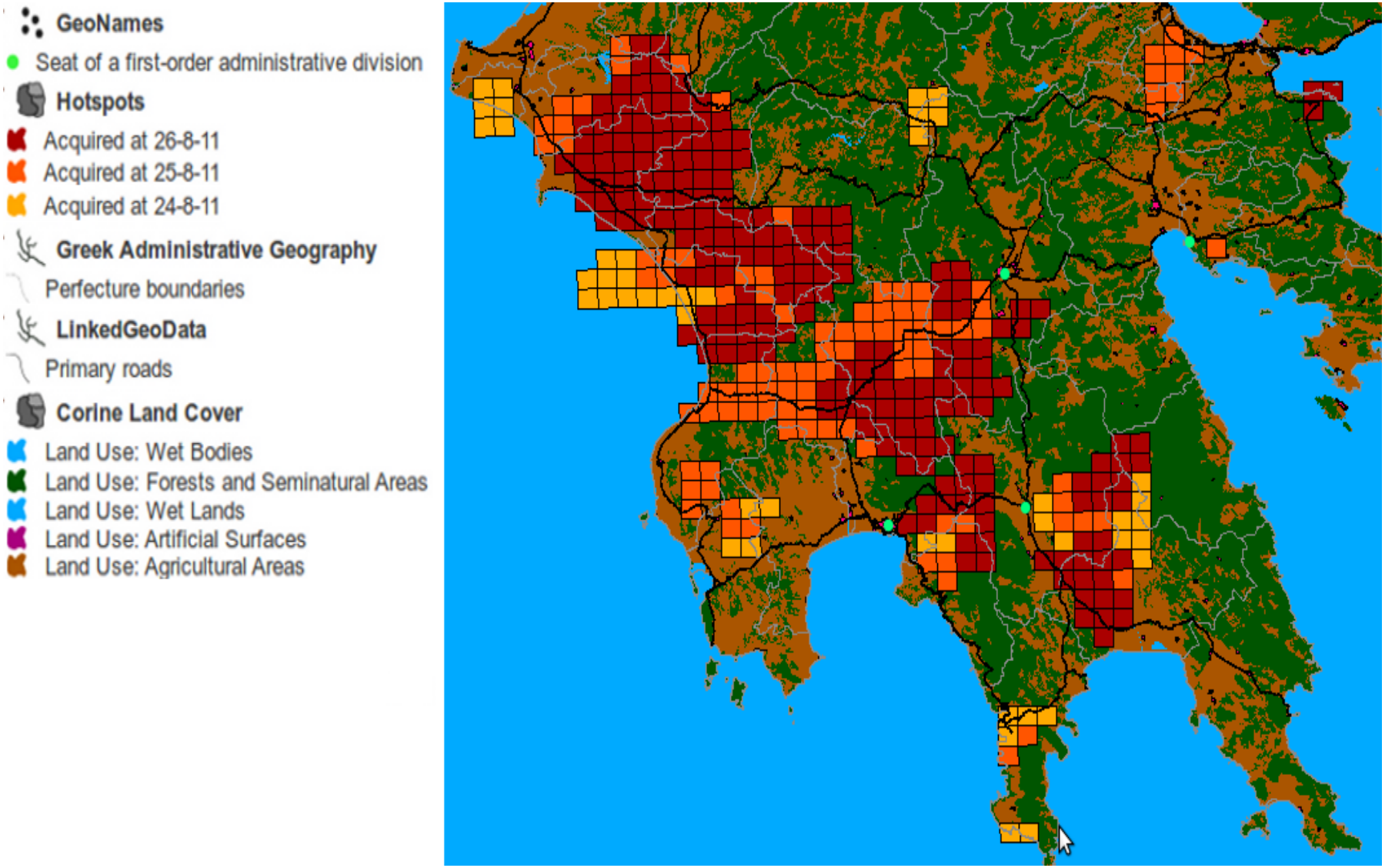
Experimental Evaluation of Strabon

- **Real linked geospatial data sets** available on the Web (DBpedia, Geonames, LGD, Pachube, SwissEx, CLC, GADM).
- **Synthetic data**
- **Geometries of various types** (points, line strings, polygons)
- Up to **1 billion triples**

Experimental Evaluation (cont'd)



Rapid mapping in the NOA Fire Monitoring Service



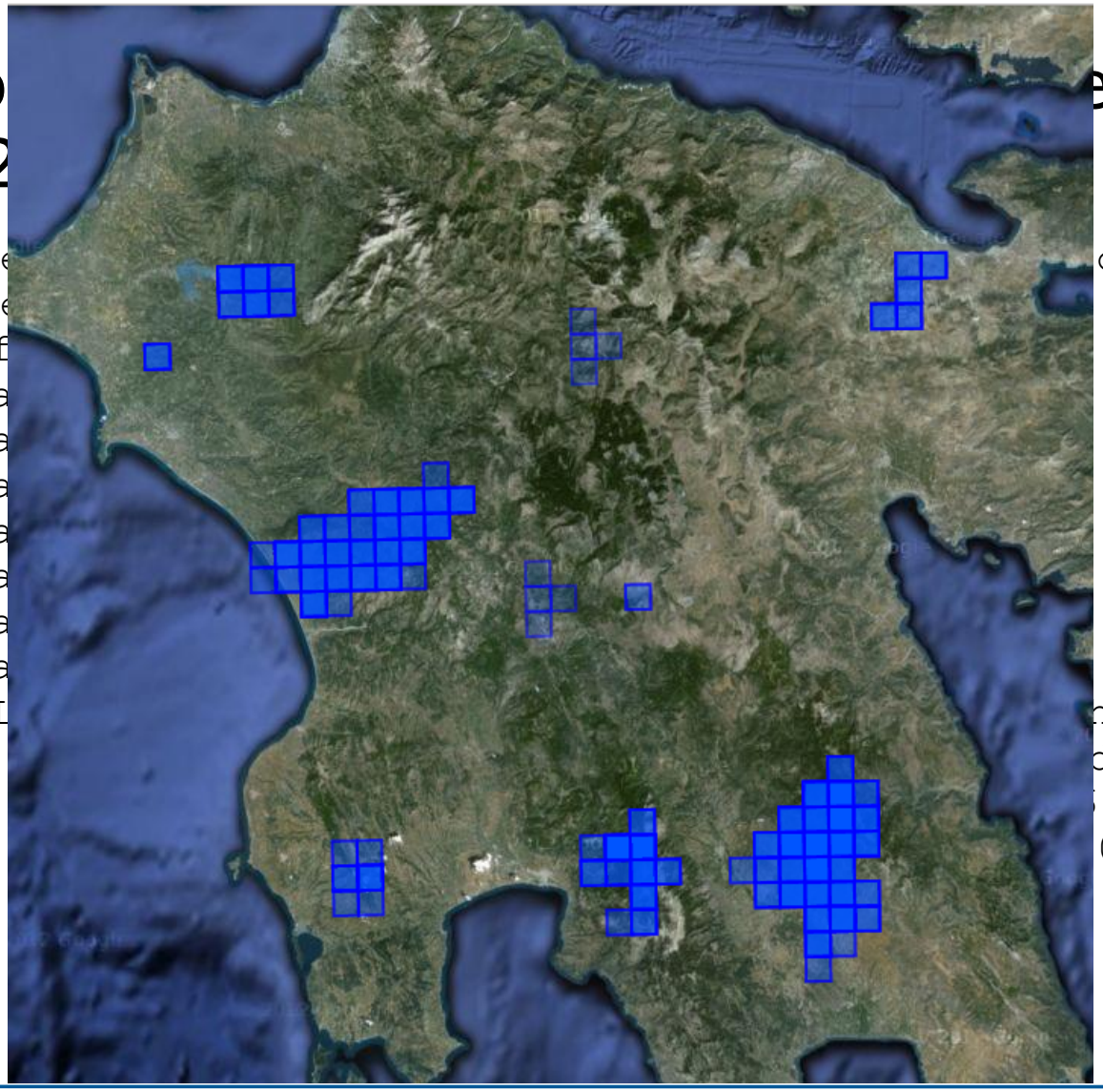
Linked Geospatial Data Currently Used

- Hotspots (MSG/SEVIRI)
- Hotspots (FIRMS/MODIS)
- Administrative regions of Greece
- Corine Land Use / Land Cover data set
- OpenStreetMap
- GeoNames

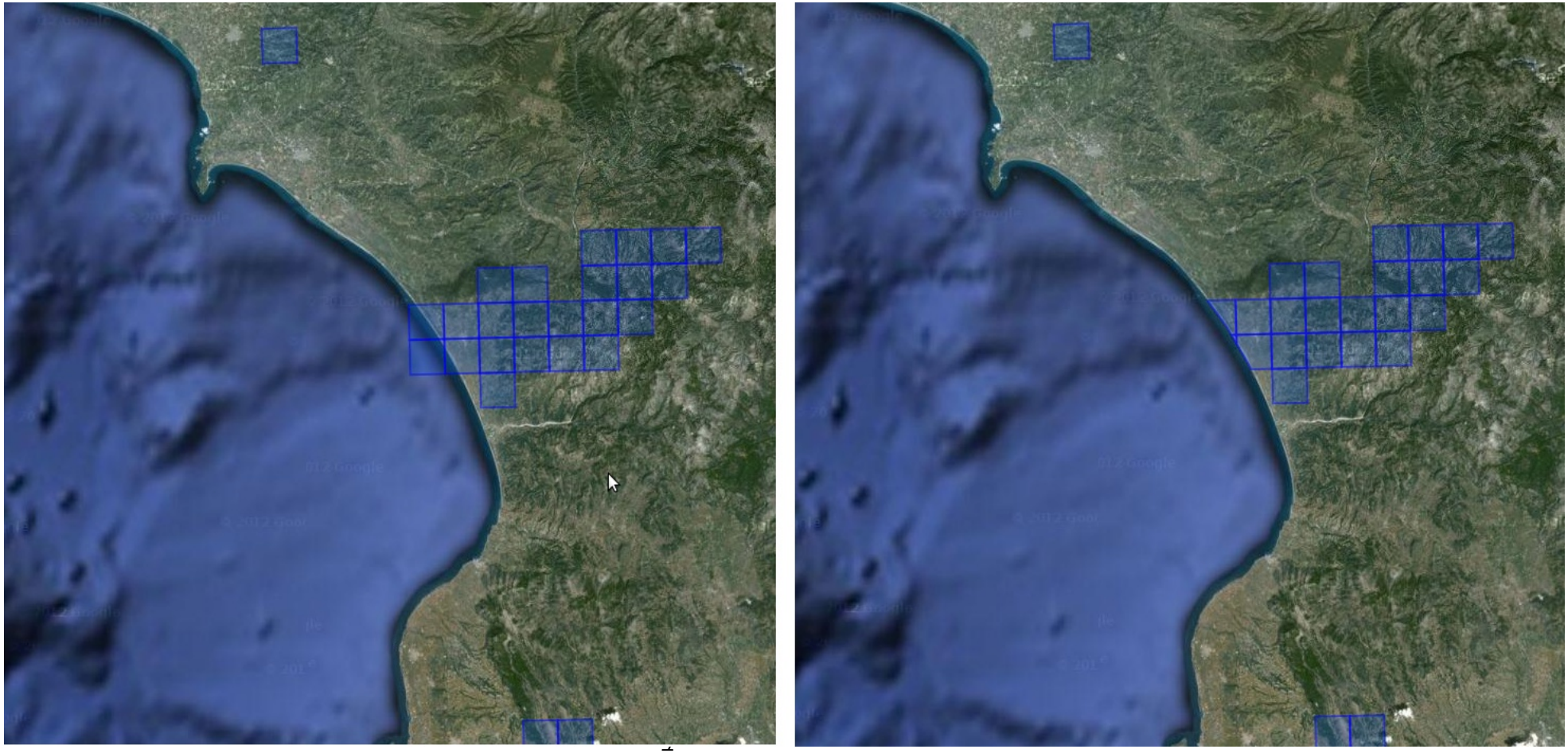
Creating a map

Get all ho ese on
24/08/2

```
SELECT ?h ?hGe ovider ?
hSensor ?hSate
WHERE {?h  rdf
          noa
          noa
          noa
          noa
          noa
          noa
          noa
          noa
          FIL                                    nAcqTime &&
                                                 dateTime).
                                                 , 23.77
                                                 027 38.36))"
```



Improve product accuracy



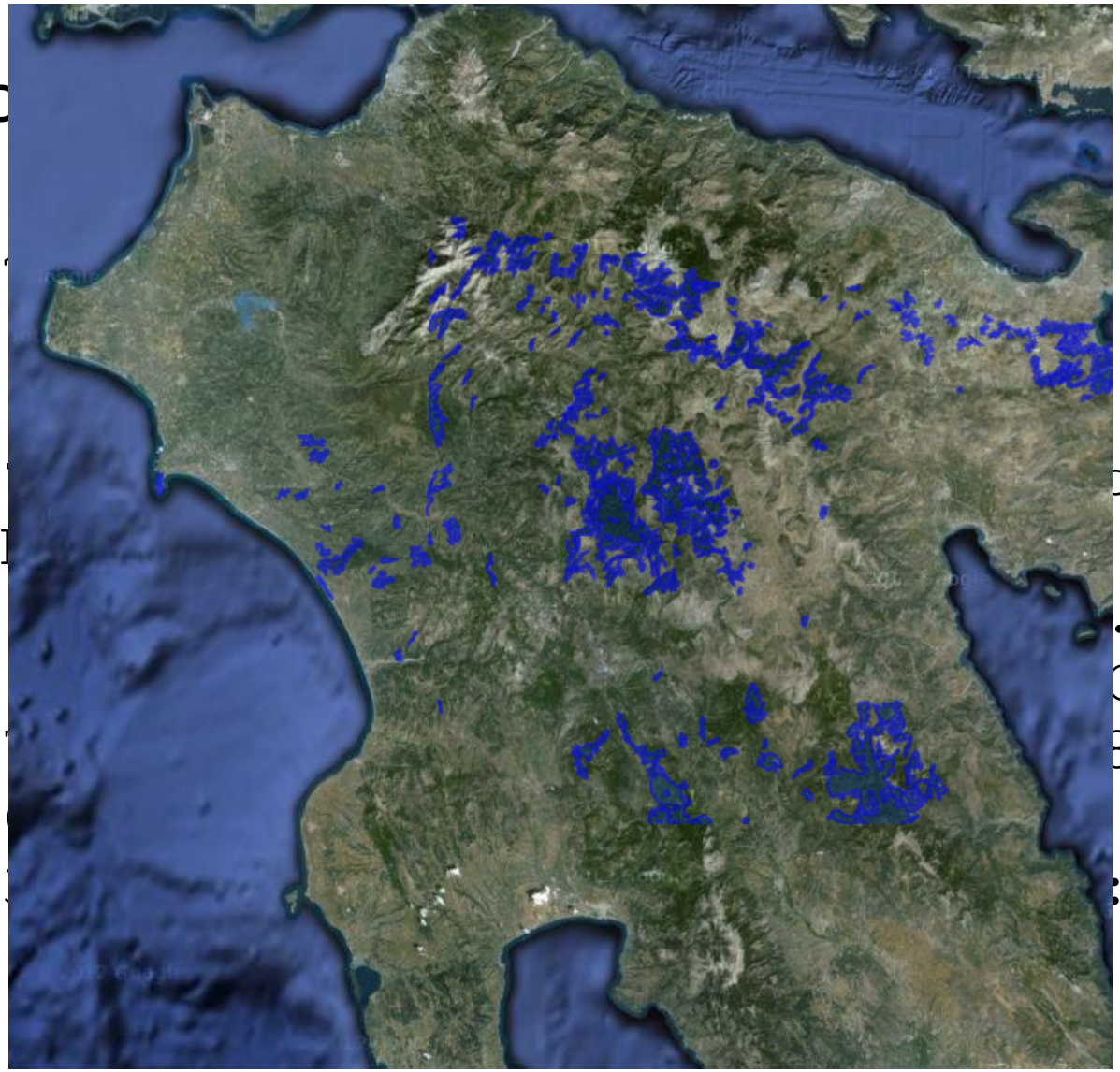
```
      FILTER( strdf:anyInteract(?hGeo, ?cGeo)}
   GROUP BY ?h ?hGeo
HAVING strdf:overlap(?hGeo, strdf:union(?cGeo))}
```

Creating a map (cont'd)

Get all c

```
SELECT
WHERE{ 1

                                                    pe.
                                                    .
                                                    ON
((21.02                                             3.36,
23.77 3                                             05,
21.027 1                                            :WKT,?
aGeo)).
}
```

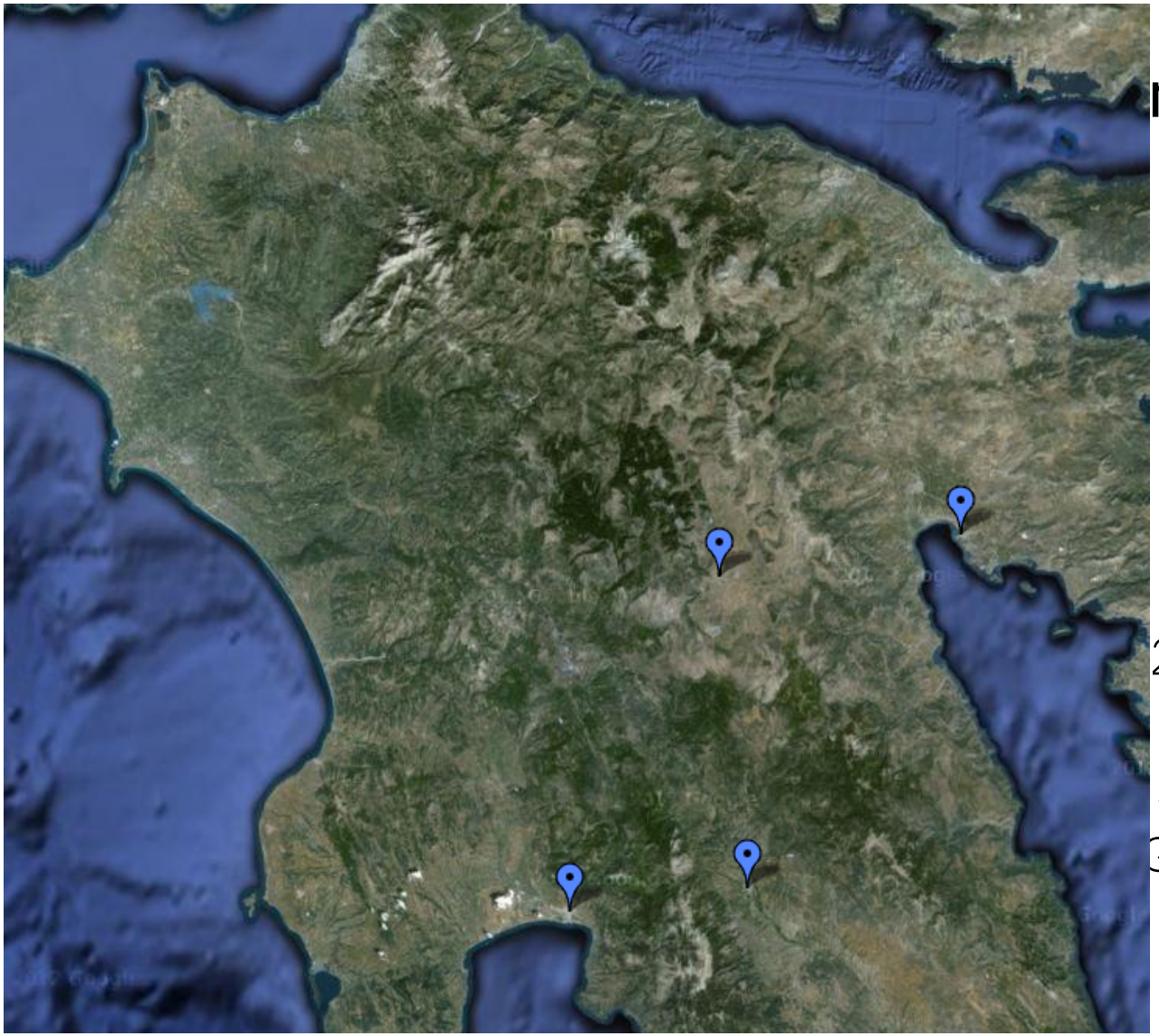


Creating a map (cont'd)

Get all
Pelop

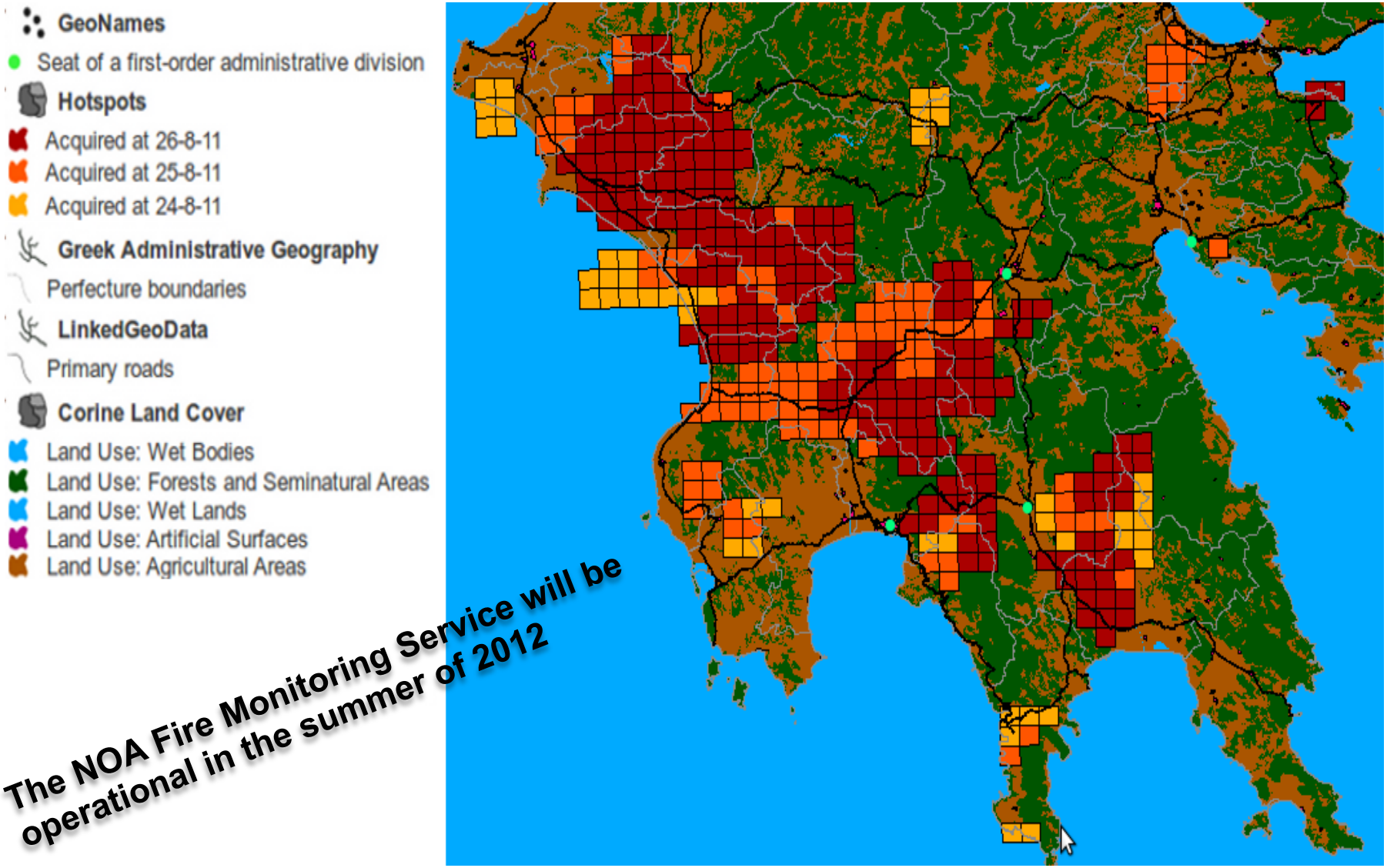
```
SELECT
WHERE{
                                                    |
                                                    21.027
                                                    77
6.05,                                               27
38.36)                                              eo) ).
}
```

Creating a map (cont'd)

Get all  nnese.

```
SELECT
WHERE{
                                                    ) .
                                                    21.51
                                                    83
                                                    51
6.41                                                Geo)).
}
```

Final map



The NOA Fire Monitoring Service will be operational in the summer of 2012

Thank you for your attention!

More information about TELEIOS can be found on
http://www.earthobservatory.eu/

Attend our tutorial on Linked Geospatial Data at