Strabon
Semantic support for EO Data Access in TELEIOS

Presenter: Kostis Kyzirakos
Outline

- Motivation
- Representing and Querying Geospatial and Temporal Information in RDF
- The Fire Monitoring Service of the National Observatory of Athens
- The TerraSAR-X Virtual Observatory of the German Aerospace Center
- Demo
- Conclusions
State of the Art in EO Data Centers

EO data center

Raw Data

EOWEB

Users

Catalogue

Processing Chains

Atmospheric → Optical → Radar

EOWEB

TELEIOS: Third User Community Workshop
Strabon: Semantic support for EO Data Access in TELEIOS
Example

- Can I pose the following query using EOWEB?

Find images taken by the SEVIRI 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 SEVIRI 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.
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**
Semantics-Based Representation and Querying of EO Data

- The data model \textit{stRDF} and the query language \textit{stSPARQL}

- The system \textit{Strabon}

\texttt{strabon.di.uoa.gr}
**RDF: Resource Description Framework**

**W3C** recommendation

RDF is a **graph data model** (+ XML syntax + semantics)

- For representing **metadata**
- For describing the semantics of information in a machine-accessible way
- 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**.

```xml
<ex:BurntArea1 noa:hasArea="23.7636"^^xsd:double/>
```
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.

- 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
stRDF: An example (1/2)
stRDF: An example (1/2)

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

```
"23.7636"^^xsd:double
```

```
"1"^^xsd:int
```

```
noa:hasArea
```

```
geo:geometry
```

```
noa:hasID
```

```
noa:BurntArea
```

```
noa:BurntArea
```

```
quiltex:BurntArea1
```

```
rdf:type
```

Spatial Data Type

Well-Known Text

Spatial Literal

(OpenGIS Simple Features)
stSPARQL: An example (1/2)

- Find all burned forests within 10kms of a city

```sparql
select ?BA ?BAGEO
where {
    ?R rdf:type noa:Region;
    geo:geometry ?Rgeo;
    noa:hasCorineLandCoverUse ?F .
    ?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)
stSPARQL: An example (1/2)
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 .
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.
?CITY rdf:type dbpedia:City ;
geo:geometry ?CGEO .
?BA 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) }
```
stSPARQL: An example (2/2)
stSPARQL: More details

- We start from **SPARQL 1.1**.

- 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.

- We add a set of temporal functions (superset of Allen’s functions) as SPARQL extension functions.

- We add appropriate geospatial and temporal extensions to **SPARQL 1.1 Update language**.
Strabon: A Scalable Geospatial RDF Store

Strabon is a scalable geospatial RDF store that supports Semantic Web technologies. It focuses on efficient processing of spatio-temporal data and provides support for various geospatial formats such as WKT and GML. Strabon uses state-of-the-art query languages like ST-RDF graphs for efficient data representation and GeoSPARQL/ST-SPARQL queries for querying spatial data.

The architecture of Strabon includes a query engine, parser, optimizer, evaluator, transaction manager, and storage manager. The storage manager can be configured with different repositories, such as SAIL and RDBMS, and with different general purpose databases like GeneralDB and monetdb. PostGIS is also integrated as a spatial database.

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
  - Producing **rapid mapping** products
The Virtual Observatory for TerraSAR-X Data

- Develop a VO that goes beyond the current EOWEB portal of DLR

- Use Semantic Web technologies for
  
  - **Representing** fire related products using ontologies
  
  - **Validating** the results of the **automatic annotation process** by correlating them with auxiliary linked geospatial data
  
  - **Assisting** the **training process** by providing the user with **contextual information** about the area of interest
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
- Datasets from Linked Open Data Cloud
  - LinkedGeoData
  - GeoNames
Linked Open Data Cloud

TELEIOS: Third User Community Workshop
Strabon: Semantic support for EO Data Access in TELEIOS
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** by correlating them with auxiliary data
- **Validate** the results of the **automatic annotation process** by correlating them with auxiliary data
- **Assist** the **training process** by providing the user with contextual information of the area of interest
Generating maps combining diverse information sources

- Generating Rapid Mapping products
- Semantic Enrichment for Hotspots
- Assisting knowledge discovery from EO images

DEMO!

Discovering EO data

Get all hotspots detected in Peloponnese on 24/08/2007.

```sparql
WHERE {
  ?h rdf:type noa:Hotspot ;
  noa:hasGeometry ?hGeo ;
  noa:hasAcquisitionTime ?hAcqTime ;
  noa:hasConfidence ?hConfidence ;
  noa:isProducedBy ?hProvider ;
  noa:hasConfirmation ?hConfirmation ;
  noa:isDerivedFromSensor ?hSensor ;
  noa:isDerivedFromSatellite ?hSatellite .
  FILTER(strdf:contains("POLYGON((21.027 38.36, 23.77 38.36, 23.77 36.05, 21.027 36.05, 21.027 38.36))"^^strdf:WKT, ?hGeo) ) .
}
```
Retrieving a map layer (1/3)

Get all coniferous forests in Peloponnese

```sql
SELECT ?a aGeo
WHERE {
  ?a rdf:type clc:Area ;
  clc:hasLandUse ?aLandUse .
  noa:hasGeometry ?aGeo .
  ?aLandUse rdf:type ?aLandUseType .
  FILTER (?aLandUseType = clc:ConiferousForest) .
}
```

Retrieving a map layer (2/3)
Retrieving a map layer (2/3)

Get all primary roads in Pelloponnese

```sparql
SELECT ?r ?rGeo
WHERE {
  ?r a ?rType ;
  noa:hasGeometry ?rGeo .
  FILTER(?rType = lgdo:Primary) .
}
```
Retrieving a map layer (3/3)

Get all capitals of prefectures of the Peloponnese.

```sql
SELECT ?feature
WHERE {
  ?feature rdf:type gn:Feature;
    noa:hasGeography ?fGeo ;
    gn:name ?fName ;
}
```
Enriching Products Using stRDF and stSPARQL

- Generating maps combining diverse information sources
- Generating Rapid Mapping products
- Semantic Enrichment for Hotspots
- Assisting knowledge discovery from EO images

DEMO!

Generating Rapid Mapping Products

ZKI (Center for Satellite Based Crisis Information)

Strabon: Semantic support for EO Data Access in TELEIOS
Generating Rapid Mapping Products: Sextant
Enriching Products Using stRDF and stSPARQL

- Generating maps combining diverse information sources
- Generating Rapid Mapping products
- Semantic Enrichment for Hotspots
- Assisting knowledge discovery from EO images

DEMO!

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. Keep land part of the polygon
  4. Eliminate false alarms in inconsistent land cover areas

- **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.
Improve the accuracy of EO data

HAVING strdf:overlap(?hGeo, strdf:union(?cGeo))

TELEIOS: Third User Community Workshop
Strabon: Semantic support for EO Data Access in TELEIOS
Enriching Products Using stRDF and stSPARQL

- Generating maps combining diverse information sources
- Generating Rapid Mapping products
- Semantic Enrichment for Hotspots
- Assisting knowledge discovery from EO images

DEMO!

Discovery Queries

Select all patches corresponding to the class **Water**
Discovery Queries

Select all patches corresponding to the class \texttt{Water} and its subclasses
Discover Correlation Between Datasets (1/2)

List the labels of all patches that are inside each Corine Land Use/Land Cover class.

```
SELECT DISTINCT ?clcLandUse ?annotation ?g
WHERE {
  ?p rdf:type dlr:Patch ;
  dlr:hasGeometry ?g ;
  dlr:hasLabel ?l .
  ?l rdf:type dlr:Label ;
  dlr:correspondsTo ?annotation .
  FILTER(strdf:anyInteract(?g, "POLYGON((…))"^^strdf:WKT)).
  ?clc rdf:type ?clcType ;
  teleios:hasCode ?clcCode ;
  teleios:hasID ?clcID ;
  teleios:hasGeometry ?clcG ;
  teleios:hasLandUse teleios:continuousUrbanFabric .
  FILTER(strdf:anyInteract(?clcG, "POLYGON((…))"^^strdf:WKT)).
  FILTER(geof:sf-contains(?clcG, ?g)) .
}
```
Discover Correlation Between Datasets (2/2)

List the labels of all OpenStreetMap classes that are inside a patch

```sql
SELECT ?annotation (GROUP_CONCAT(DISTINCT ?lgdType) AS ?lgdTypes)
WHERE {
  ?p rdf:type dlr:Patch ;
  dlr:hasGeometry ?g ;
  dlr:hasLabel ?l .
  ?l rdf:type dlr:Label ;
  dlr:correspondsTo ?annotation .
  FILTER (strdf:anyInteract (?g, "POLYGON((...))"^^strdf:WKT)).
  ?lgd lgdont:directType ?lgdDirectType .
  OPTIONAL { ?lgd a ?lgdType .
    FILTER((?lgdType != lgdont:Node) && (?lgdType != ?lgdDirectType))}.}
  FILTER (strdf:anyInteract (?lgdGeo, "POLYGON((...))"^^strdf:WKT)) .
  FILTER (geof:sf-contains (?g, ?lgdGeo)) .
}
GROUP BY ?annotation
ORDER BY ?annotation
```
Representing and Querying Topological Information

And now for something completely different...
Hotspot example (reminder)

part of West Greece

Olympia
Hotspot example (reminder)

RDF representation

```plaintext
gag:WestGreece rdf:type geo:Feature.
gag:WestGreece strdf:hasGeometry
   "POLYGON (( ... ))"^^strdf:WKT.

gag:Olympia rdf:type geo:Feature.
gag:Olympia strdf:hasGeometry
   "POLYGON (( ... ))"^^strdf:WKT.

noa:hotspot1 rdf:type noa:Hotspot.
noa:hotspot1 strdf:hasGeometry
   "POLYGON (( ... ))"^^strdf:WKT.
```
Find fires **inside** West Greece

\[
\text{SELECT } \ ?f \\
\text{WHERE } \{ \\
\quad ?h \text{ rdf:type noa:Hotspot.} \\
\quad ?h \text{ noa:correspondsTo } \ ?f. \\
\quad ?f \text{ rdf:type noa:Fire.} \\
\quad ?h \text{ strdf:hasGeometry } \ ?hGeo. \\
\quad \text{gag:WestGreece strdf:hasGeometry } \ ?wgGeo. \\
\quad \text{FILTER } \left( \text{strdf:contains} \left( ?wgGeo, \ ?hGeo \right) \right) \\
\}
\]
Hotspot example (revisited)

part of West Greece

Olympia
Hotspot example (revisited)
RDF representation

```
gag:WestGreece rdf:type geo:Feature.
gag:WestGreece strdf:hasGeometry
  "POLYGON (( ... ))"^^strdf:WKT.

gag:Olympia rdf:type geo:Feature.
gag:Olympia strdf:hasGeometry
  "POLYGON (( ... ))"^^strdf:WKT.

noa:hotspot1 rdf:type noa:Hotspot.
noa:hotspot1 strdf:hasGeometry
  "POLYGON (( ... ))"^^strdf:WKT.

noa:fire rdf:type noa:Fire.
noa:hotspot1 geo:sfContains noa:fire.
```

Topology vocabulary extension of GeoSPARQL
Find fires **inside** West Greece

```sparql
SELECT ?f
WHERE {
  ?h rdf:type noa:Hotspot.
  ?h noa:correspondsTo ?f.
  ?h strdf:hasGeometry ?hGeo.
  FILTER (strdf:contains(?wgGeo, ?hGeo))
}

The geometry of fires is not available!
```
Querying Topological Information

Find fires inside West Greece

```
SELECT ?f
WHERE {
}
```

**Beyond** stSPARQL and the topology vocabulary extension of GeoSPARQL
Conclusions

- We developed the data model **stRDF** and the query language **stSPARQL** for representing and querying **geospatial data** that may **change over time**
- The Fire Monitoring Service of the National Observatory of Athens
- The TerraSAR-X Virtual Observatory of the German Aerospace Center
Discussion

- Use **higher-level languages**, stop worrying about how to store and manage metadata, just **focus** on the actual **processing**.

- Express common earth observation operations easily using the **stSPARQL/GeoSPARQL** queries instead of using a lengthy **C** program.

- **Rapid prototyping** without the need to recompile everything.

- **Integration** with publically available **linked open geospatial data**.
Thank you for your attention!

Ερωτήσεις;

- **Strabon** [http://strabon.di.uoa.gr](http://strabon.di.uoa.gr)
  Manolis Koubarakis, Kostis Kyzirakos, Manos Karpathiotakis, Charalampos Nikolaou, Giorgos Garbis, Konstantina Bereta, Kallirroi Dogani, Stella Giannakopoulou and Panayiotis Smeros.
  - Mercurial repository: [http://hg.strabon.di.uoa.gr](http://hg.strabon.di.uoa.gr)