Representing and Querying Linked Geospatial Data

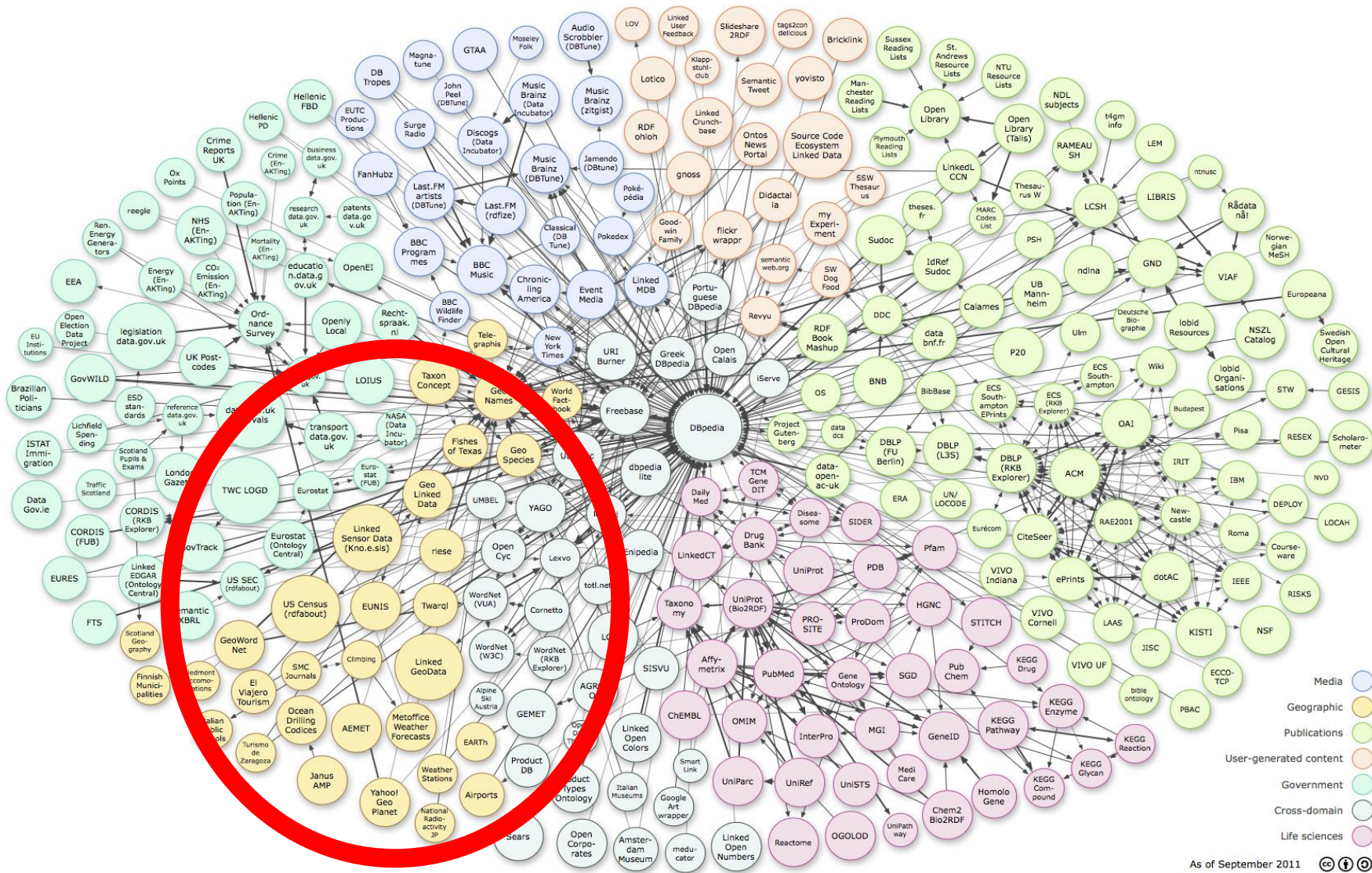## Kostis Kyzirakos
kostis@cwi.nl

CWI

Centrum voor Wiskunde en Informatica
Database Architectures group

University of Athens
School of Science
Faculty of Informatics and Telecommunications

Linked Open Data Cloud

As of September 2011

RDF: Resource Description Framework

W3C recommendation

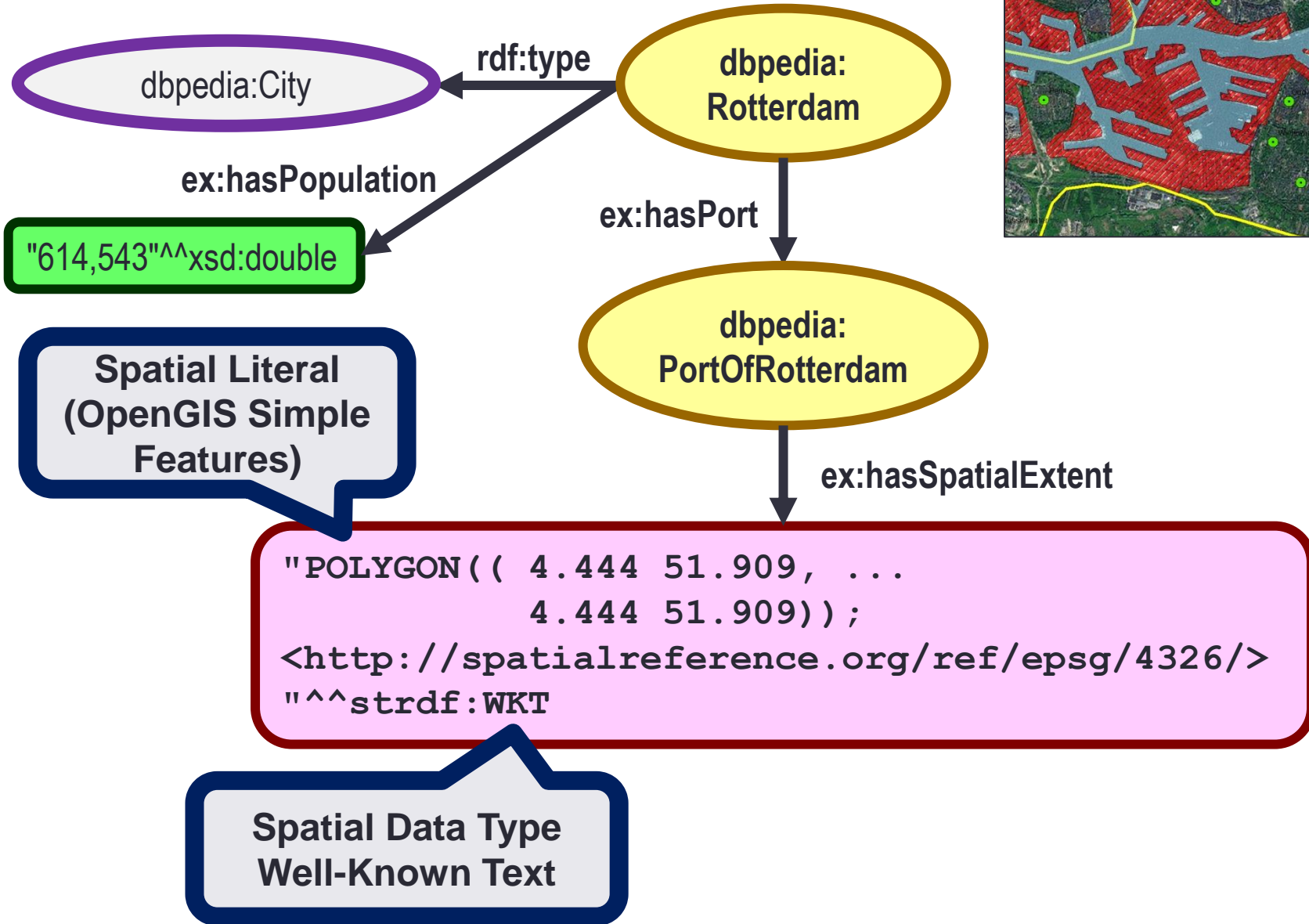RDF is a **graph data model** (+ XML syntax + semantics)

- 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**

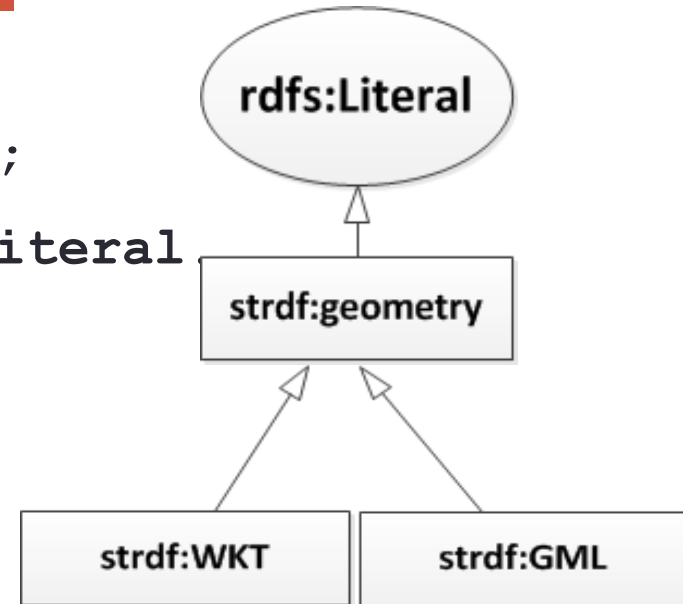- Statements are represented as **triples**, consisting of a **subject**, **predicate** and **object**.

dbpedia:
Rotterdam

ex:hasPopulation →

" 614,543"^^xsd:integer

Rotterdam

Rotterdam

dbpedia:City

rdf:type

dbpedia: Rotterdam

ex:hasPopulation

"614,543"^^xsd:double

Spatial Literal (OpenGIS Simple Features)

ex:hasPort

dbpedia: PortOfRotterdam

ex:hasSpatialExtent

```
"POLYGON(( 4.444 51.909, ...
          4.444 51.909));
<http://spatialreference.org/ref/epsg/4326/>
"^^strdf:WKT
```

Spatial Data Type Well-Known Text

The stRDF Data Model
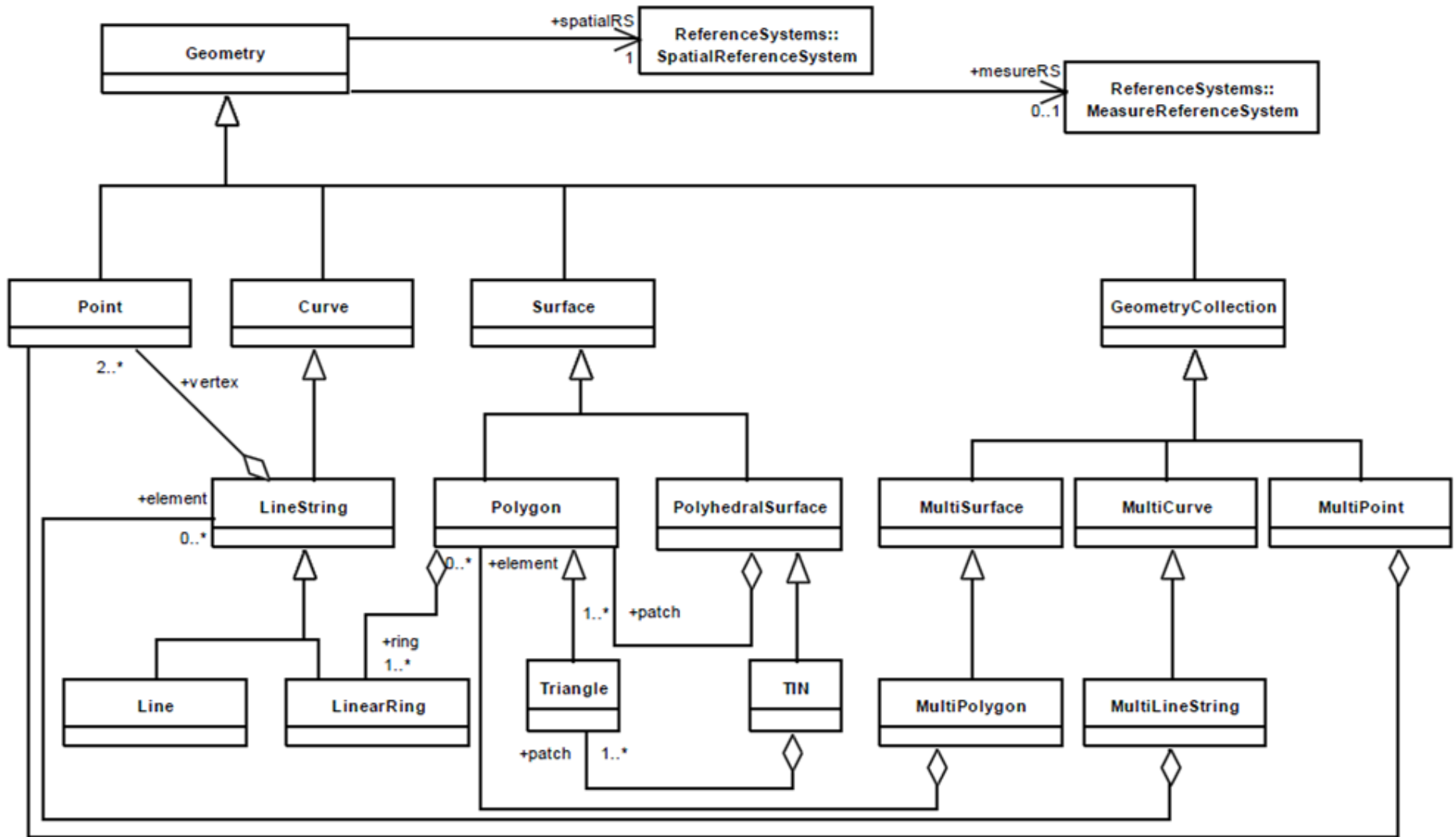

strdf:geometry rdf:type rdfs:Datatype;

        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 industrial areas of the port of Rotterdam

stSPARQL: An example (1/2)

Find all industrial areas of the port of Rotterdam
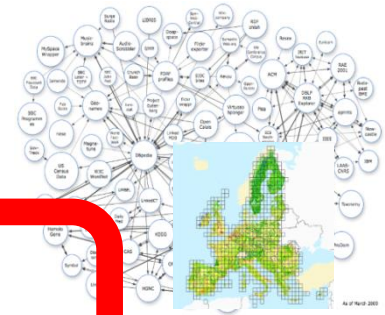
```
SELECT ?area
WHERE {
   ?port   rdf:type dbpedia:Port ;
           clc:hasLandCover clc:Port;
           ex:hasSpatialExtent ?portGeom .

   ?area   rdf:type osm:Industrial;
           ex:hasSpatialExtent ?aGeom.

   FILTER(strdf:intersects(?portGeom, ?aGeom))
}
```

stSPARQL: An example (1/2)

Find all industrial areas of the port of Rotterdam

```
SELECT ?area
WHERE {

    ?port    rdf:type dbpedia:Port ;
             clc:hasLandCover clc:Port;
             ex:hasSpatialExtent ?portGeom .

    ?area    rdf:type osm:Industrial;
             ex:hasSpatialExtent ?aGeom.

    FILTER(strdf:intersects(?portGeom, ?aGeom))
}
```

Spatial Function

stSPARQL: An example (2/2)

Isolate the industrial areas that lie in the port of Rotterdam

stSPARQL: An example (2/2)

Isolate the industrial areas that lie in the port of Rotterdam

```
SELECT
 (strdf:intersection(?portGeom ,
                     strdf:union(?aGeom))
 AS ?industrialPart)
WHERE {
   ?port   rdf:type dbpedia:Port ;
           clc:hasLandCover clc:Port;
           ex:hasGeometry ?portGeom .

   ?area   rdf:type osm:Industrial;
           ex:hasSpatialExtent ?aGeom.

   FILTER(strdf:intersects(?portGeom, ?aGeom))
}
GROUP BY ?portGeom
```
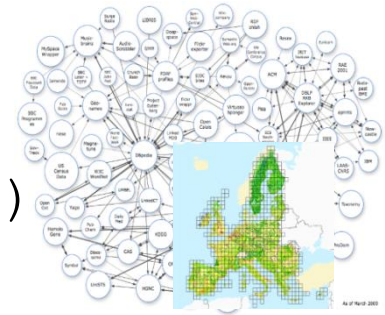
stSPARQL: An example (2/2)

Isolate the industrial areas that lie in the port of Rotterdam

Spatial Aggregate

```
SELECT
(strdf:intersection ?portGeom ,
                    strdf:union ?aGeom
AS ?industrialPart)
WHERE {
    ?port   rdf:type dbpedia:Port ;
            clc:hasLandCover clc:Port;
            ex:hasGeometry ?portGeom .

    ?area   rdf:type osm:Industrial;
            ex:hasSpatialExtent ?aGeom.

    FILTER(strdf:intersects ?portGeom, ?aGeom)
}
GROUP BY ?portGeom
```
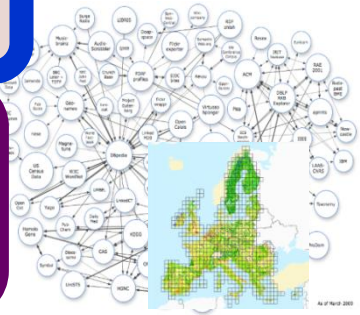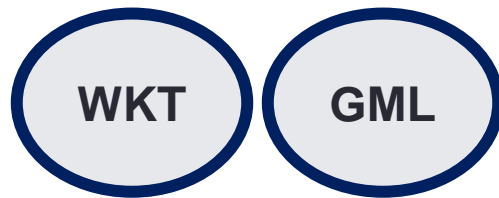
stSPARQL: Geospatial SPARQL 1.1

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`)

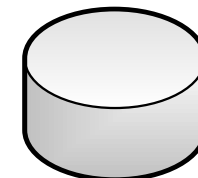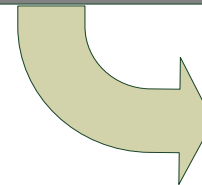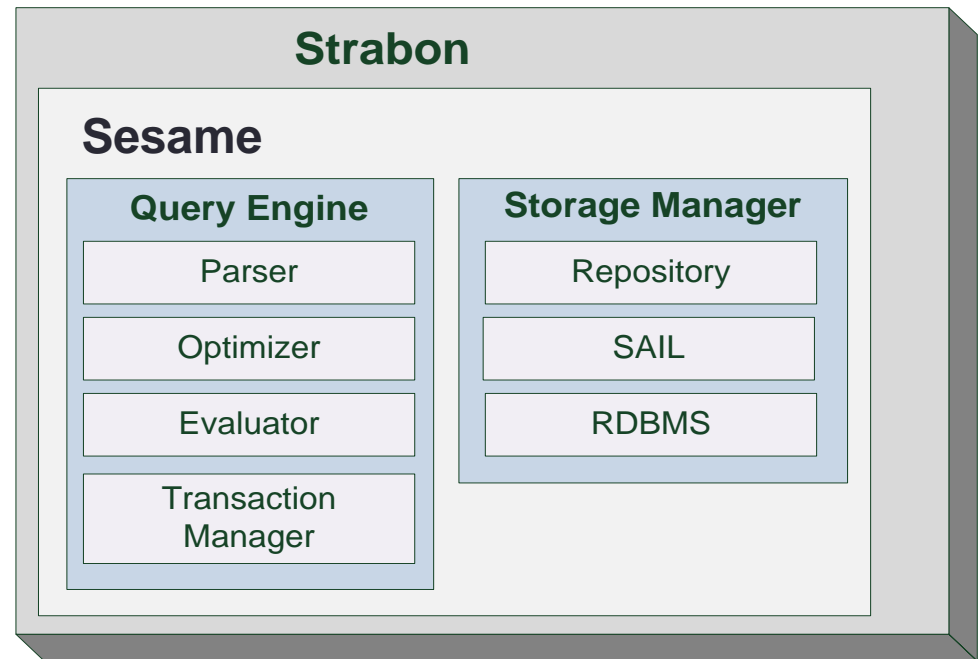- **Spatial aggregate** functions (e.g., `strdf:union`, `strdf:extent`)
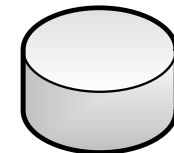
The system Strabon

WKT   **GML**

stRDF graphs

stSPARQL/ GeoSPARQL queries

Strabon

Sesame

Query Engine

Parser

Optimizer

Evaluator

Transaction Manager

Storage Manager

Repository

SAIL

RDBMS

GeneralDB

PostGIS

monet db

http://strabon.di.uoa.gr

Real-world Workload: 100 million triples – cold caches

Thematic selectivity: 0.1%

Thematic selectivity: 100%

Advertisements

- Geospatial RDF store Strabon
  http://strabon.di.uoa.gr/

- Visualization and exploration of time-evolving linked geospatial data
  http://sextant.di.uoa.gr/

- Geographica: A Benchmark for Geospatial RDF Stores
  http://geographica.di.uoa.gr/

[ESWC '10,
ISWC '12]

[ESWC '13,
ISWC '13]

[ISWC '13]

Questions?

- Geospatial RDF store Strabon
  http://strabon.di.uoa.gr/

- Visualization and exploration of time-evolving linked geospatial data
  http://sextant.di.uoa.gr/

- Geographica: A Benchmark for Geospatial RDF Stores
  http://geographica.di.uoa.gr/

[ESWC '10, ISWC '12]

[ESWC '13, ISWC '13]

[ISWC '13]

Kostis Kyzirakos
`kostis@cwi.nl`

CWI monet*db*