Reinforcing the AI4EU Platform by Advancing
Earth Observation Intelligence, Innovation & Adoption

Linked Data Tools in AI4Copernicus

Despina-Athanasia Pantazi
National Kapodistrian University of Athens (UoA)

BiDS, November 06 2023

Why Linked Data?

The vision of **linked data** is to go from a Web of documents to a Web of data:

- Unlock open data dormant in their silos

- Make it available on the Web using Semantic Web technologies (HTTP, URIs, RDF, SPARQL)

- Interlink it with other data (e.g., from the European data portal)

SPARQL and GeoSPARQL

- SPARQL is the standard Query Language for RDF and a W3C Recommendation. It allows for a query to consist of triple patterns, conjunctions, disjunctions, and optional patterns.
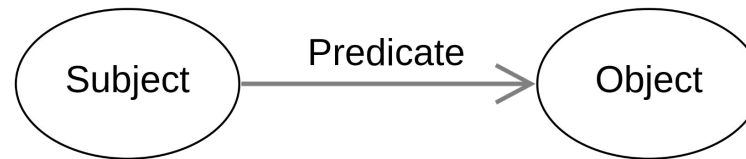
- GeoSPARQL is a standard for representation and querying of geospatial linked data for the Semantic Web from the Open Geospatial Consortium (OGC).

Query Example in GeoSPARQL

Find all potato fields in "Kirchberg" area that are less than 2km away from Elbe river.

```
SELECT DISTINCT ?field
WHERE
{
  ?field fso:hasLabelName "Potato"^^xsd:string.
  ?field geo:hasGeometry ?fieldGeometry.
  ?fieldGeometry geo:asWKT ?fieldWKT.

  ?aoi rdf:type gadm:AdministrativeUnit3.
  ?aoi gadm:has_NAME_3 "Kirchberg".
  ?aoi geo:hasGeometry ?aoiGeometry.
  ?aoiGeometry geo:asWKT ?aoiWKT.

  FILTER (geof:sfContains(?aoiWKT, ?fieldWKT))

  ?river hydro:hasName "Elbe"^^xsd:string.
  ?river hydro:hasNode ?riverPart.
  ?riverPart geo:hasGeometry ?riverPartGeometry.
  ?riverPartGeometry geo:asWKT ?riverPartWKT.

  FILTER (geof:distance(?fieldWKT, ?riverPartWKT, uom:metre) < 2000)
}
```
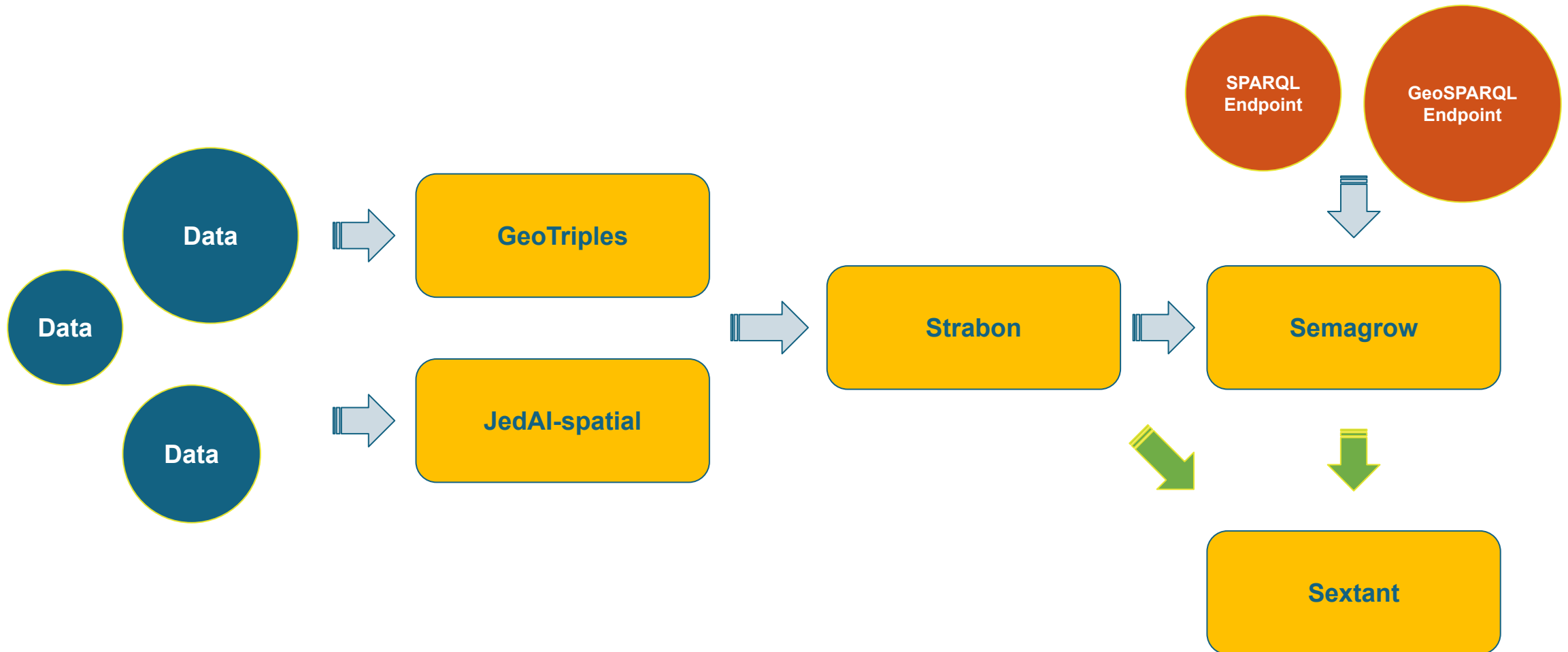
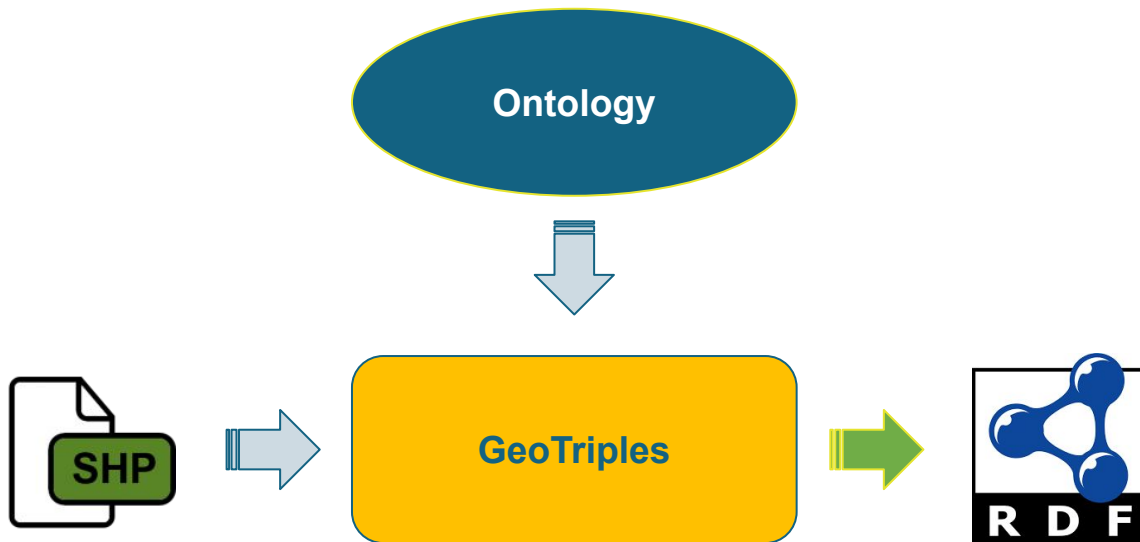# Linked Data Tools in AI4Copernicus

Linked Data Pipeline

GeoTriples

We use GeoTriples to transform the data into the RDF format, utilizing an ontology.


```
@prefix fso: <http://ai.di.uoa.gr/fs/ontology/>.
@prefix fsr: <http://ai.di.uoa.gr/fs/resource/>.
@prefix geo: <http://www.opengis.net/ont/geosparql#>.

map:DOURO_precipitation_09_2021_geometry
    rr:logicalTable [ rr:tableName "`DOURO_precipitation_09_2021`"; ];

    rr:subjectMap [
        rr:template "http://ai.di.uoa.gr/fs/resource/PREC_PR_M9_2021{`gid`}_Geometry";

    rr:predicateObjectMap [
        rr:predicate ogc:asWKT;
        rr:objectMap [
            rr:datatype ogc:wktLiteral;
            rrx:function rrxf:asWKT;
            rrx:argumentMap
            (
                [ rr:column "`the_geom`"; ]
            )
        ];
    ];
    .

map:DOURO_precipitation_09_2021
    rr:logicalTable [ rr:tableName "`DOURO_precipitation_09_2021`"; ];
    rr:subjectMap [
        rr:class fso:FoodSecurityObservation; #rdf:type fso:FoodSecurityObservation
        rr:template "http://ai.di.uoa.gr/fs/resource/FSObservation_PR_M9_2021{`gid`}";

    rr:predicateObjectMap [    #fso:hasStartDate
        rr:predicate fso:hasStartDate;
        rr:objectMap [
            rr:datatype xsd:dateTime;
            rr:template "2021-09-01T00:00:00";
        ];
    ];

    rr:predicateObjectMap [    #fso:hasEndDate
        rr:predicate fso:hasEndDate;
        rr:objectMap [
            rr:datatype xsd:dateTime;
            rr:template "2021-09-30T00:00:00";
        ];
    ];

map:DOURO_precipitation_09_2021_WaterAvailability
    rr:logicalTable [ rr:tableName "`DOURO_precipitation_09_2021`"; ];
```
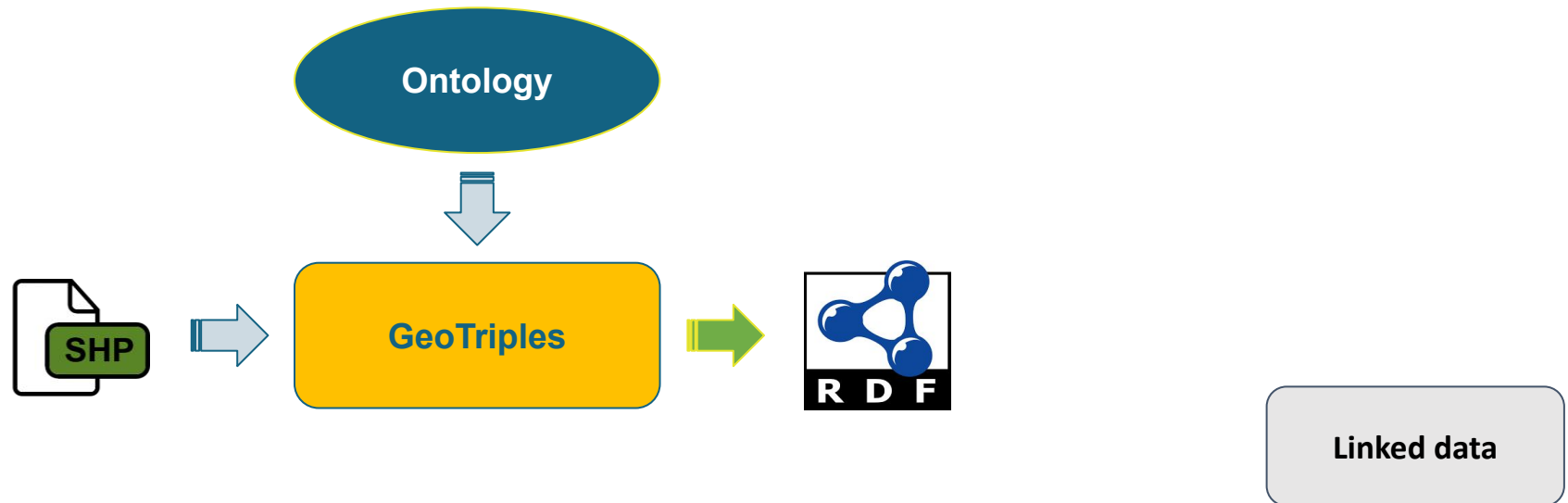
Mapping file

GeoTriples

We use GeoTriples to transform the data into the RDF format, utilizing an ontology.

Ontology → **GeoTriples** → **RDF**

SHP → GeoTriples → RDF

Linked data

```
<http://ai.di.uoa.gr/fs/resource/FSObservation_PR_M4_20211> <http://ai.di.uoa.gr/fs/ontology/hasStartDate> "2021-04-01T00:00:00"^^<http://www.w3.org/2001/XMLSchema#dateTime> .
<http://ai.di.uoa.gr/fs/resource/FSObservation_PR_M4_20212> <http://ai.di.uoa.gr/fs/ontology/hasStartDate> "2021-04-01T00:00:00"^^<http://www.w3.org/2001/XMLSchema#dateTime> .
<http://ai.di.uoa.gr/fs/resource/FSObservation_PR_M4_20213> <http://ai.di.uoa.gr/fs/ontology/hasStartDate> "2021-04-01T00:00:00"^^<http://www.w3.org/2001/XMLSchema#dateTime> .
<http://ai.di.uoa.gr/fs/resource/FSObservation_PR_M4_20211> <http://ai.di.uoa.gr/fs/ontology/hasEndDate> "2021-04-30T00:00:00"^^<http://www.w3.org/2001/XMLSchema#dateTime> .
<http://ai.di.uoa.gr/fs/resource/FSObservation_PR_M4_20212> <http://ai.di.uoa.gr/fs/ontology/hasEndDate> "2021-04-30T00:00:00"^^<http://www.w3.org/2001/XMLSchema#dateTime> .
<http://ai.di.uoa.gr/fs/resource/FSObservation_PR_M4_20213> <http://ai.di.uoa.gr/fs/ontology/hasEndDate> "2021-04-30T00:00:00"^^<http://www.w3.org/2001/XMLSchema#dateTime> .
```

JedAI-spatial

We use JedAI-spatial to detect spatial links between datasets. The resulting triples, allow us to materialize the spatial intersections in these datasets, which lowers the query execution times for those queries that utilize these links.
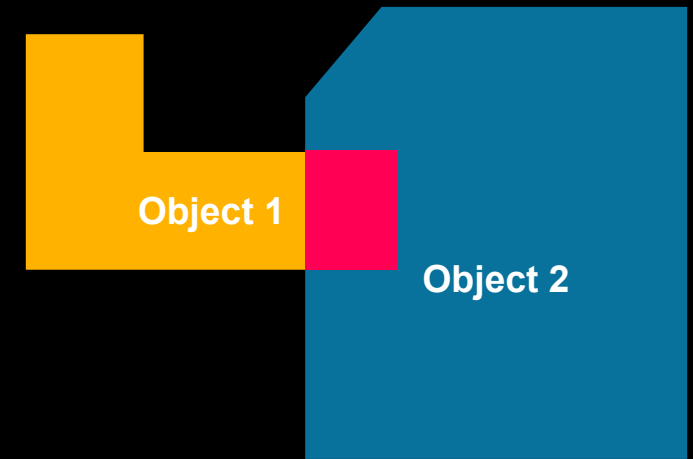
<Object1> ex:intersects <Object2> .

```
SELECT ?x                          1
WHERE {
      ?x ex:intersects ex:Athens .
}
```
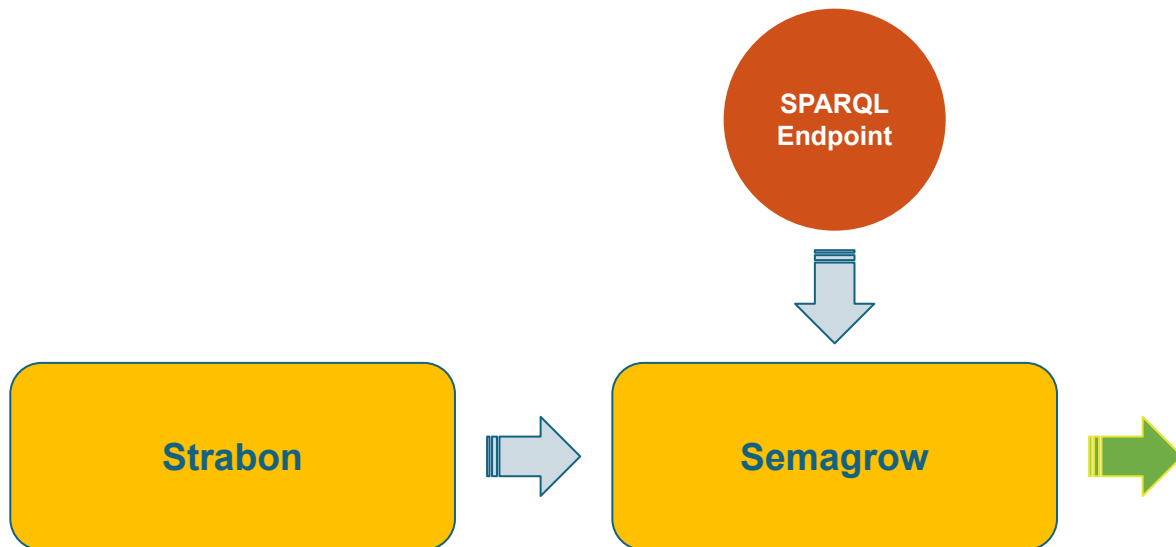
```
SELECT ?x                          2
WHERE {
      ?x geo:asWKT ?wkt .
      ex:Athens geo:asWKT ?wktA .

      FILTER (geof:sfIntersects (?wkt, ?wktA))
}
```

Strabon and Semagrow

Once the data is in the RDF format, we use the system Strabon to store them. Semagrow can federate Strabon with external SPARQL endpoints to answer queries that involve more data sources.

Find all potato fields in "Kirchberg" area that are less than 2km away from Elbe river.

```
SELECT DISTINCT ?field
WHERE
{
   ?field fso:hasLabelName "Potato"^^xsd:string.
   ?field geo:hasGeometry ?fieldGeometry.
   ?fieldGeometry geo:asWKT ?fieldWKT.

   ?aoi rdf:type gadm:AdministrativeUnit3.
   ?aoi gadm:has_NAME_3 "Kirchberg".
   ?aoi geo:hasGeometry ?aoiGeometry.
   ?aoiGeometry geo:asWKT ?aoiWKT.

   FILTER (geof:sfContains(?aoiWKT, ?fieldWKT))

   ?river hydro:hasName "Elbe"^^xsd:string.
   ?river hydro:hasNode ?riverPart.
   ?riverPart geo:hasGeometry ?riverPartGeometry.
   ?riverPartGeometry geo:asWKT ?riverPartWKT.

   FILTER (geof:distance(?fieldWKT, ?riverPartWKT,
   uom:metre) < 2000)
}
```
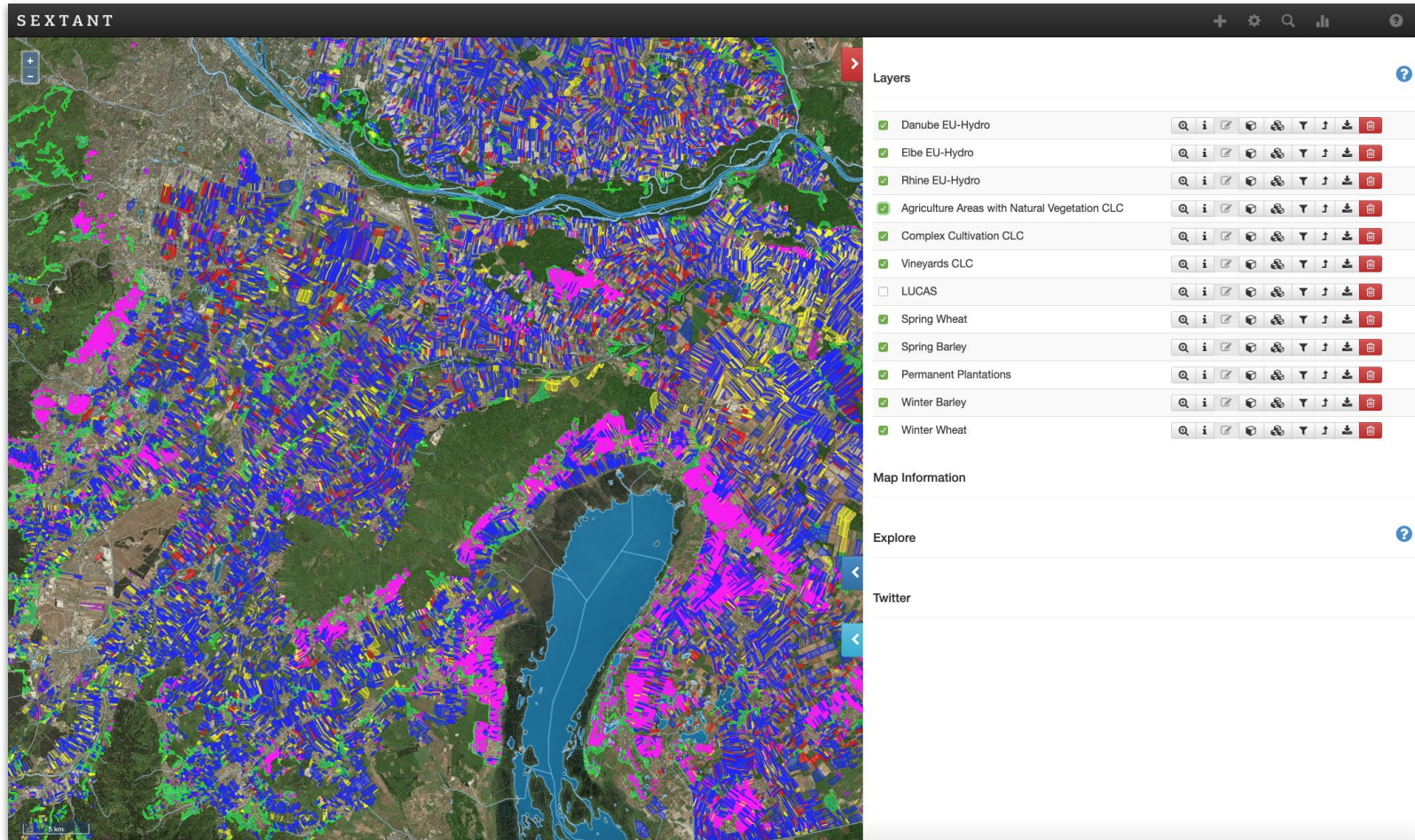
Visualize queries in Sextant

Thank You!

Any Questions?