Linked Data 101

for geospatial

Paul van Genuchten ISRIC.org
Marco Neumann KONA

FOSS4G
BUENOS AIRES 2021

29 September 2021, Ushuaia

Table of Contents

- Context
- Linked data 101
- Use Cases
- GeoSPARQL

FOSS4G
BUENOS AIRES 2021

Context

Some history of the authors

Paul van Genuchten

- ISRIC; World Soil Information
- SDI & Standardisation for >10yr
- GeoNetwork & pygeoapi

FOSS4G
BUENOS AIRES 2021

Marco Neumann

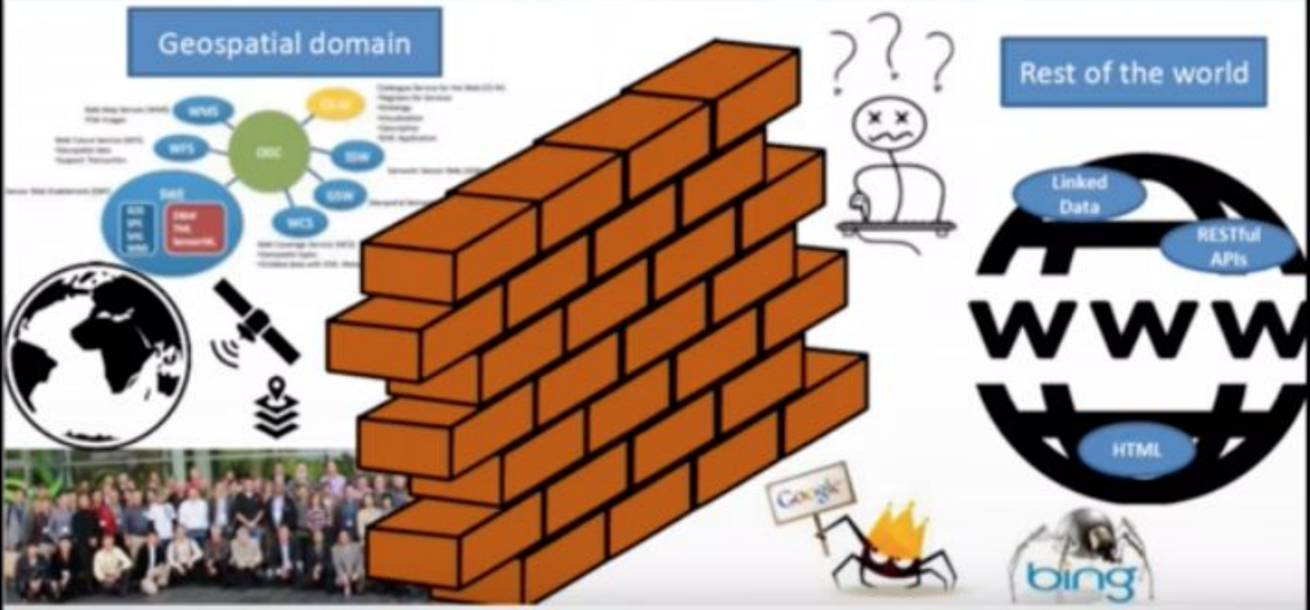
- Information Scientist KONA
- Semantic Technology Expert
- Creator of GeoSPARQL.org (fuseki)

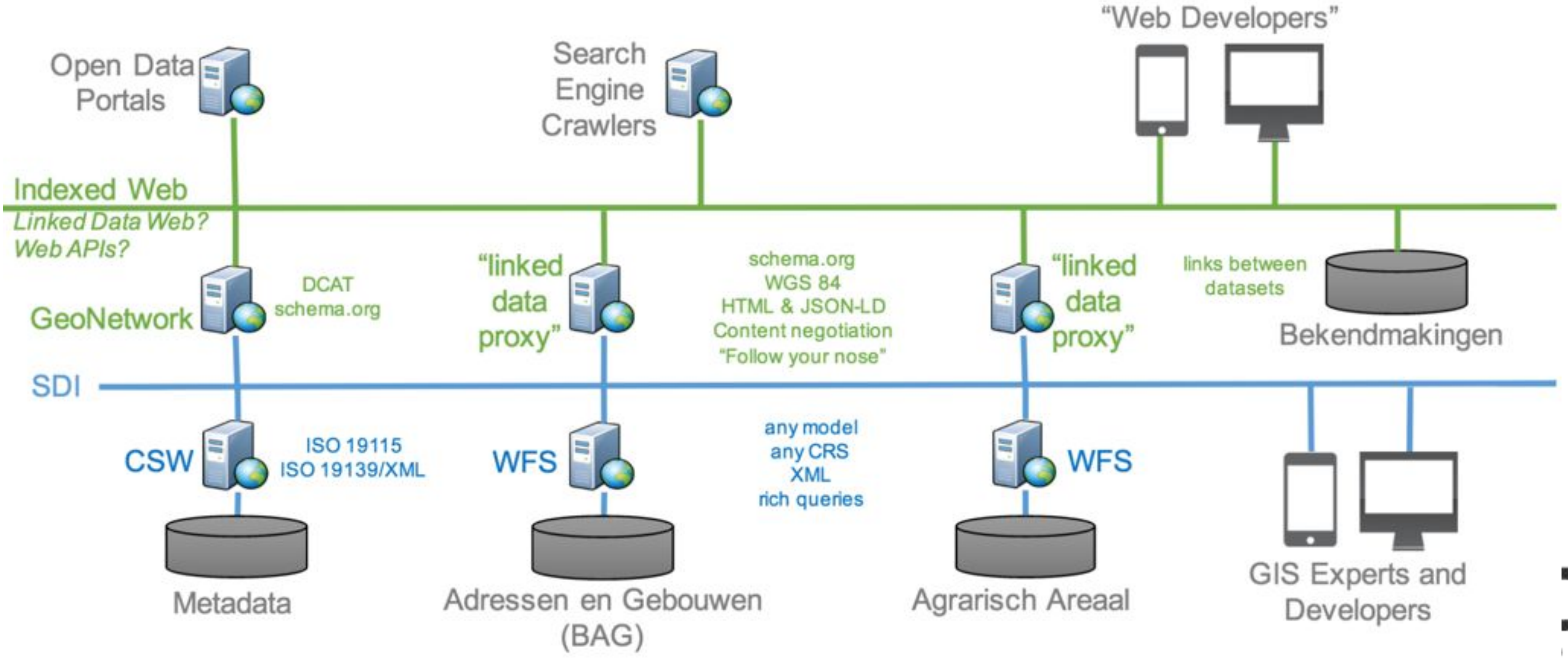
FOSS4G

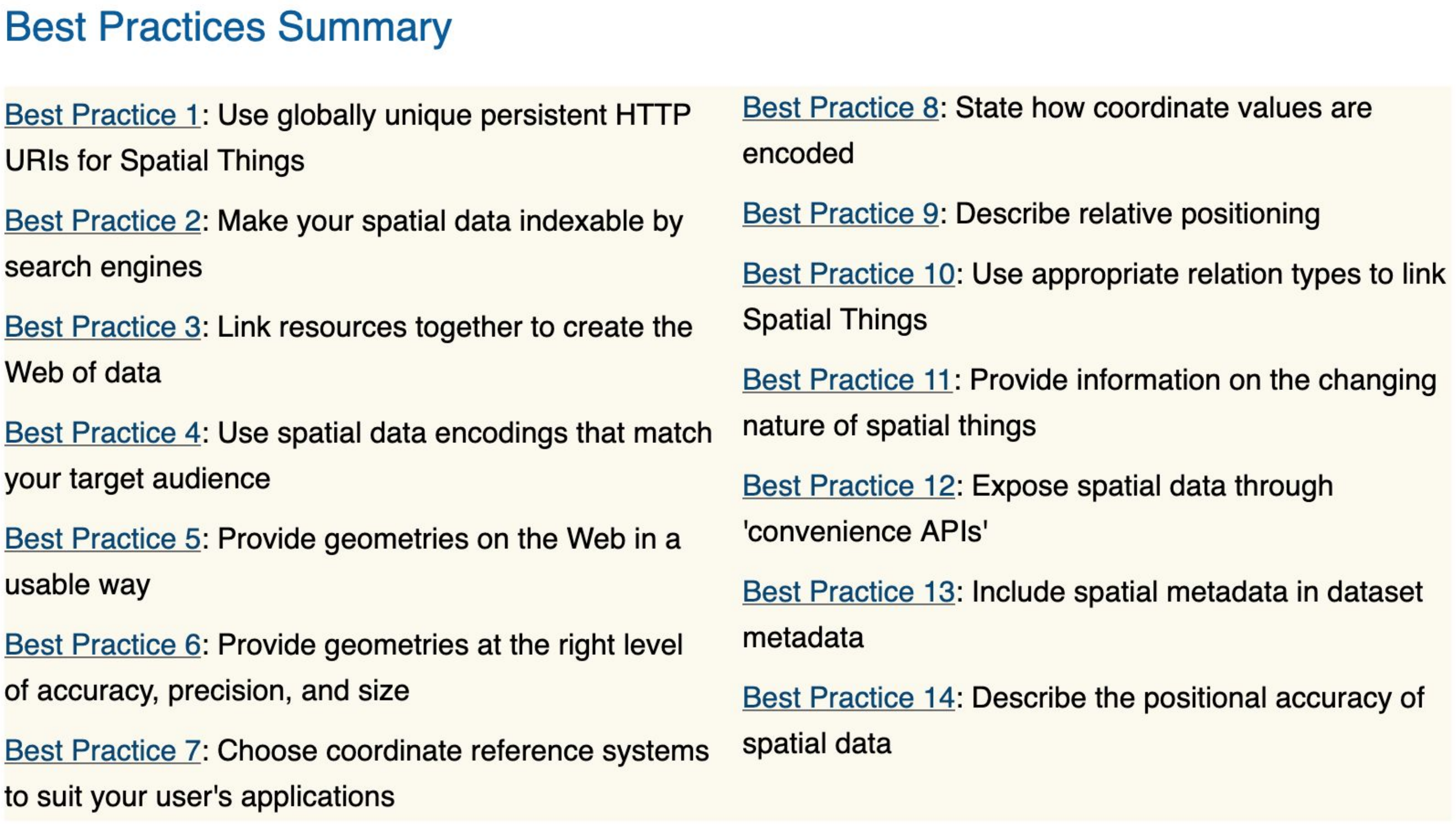
BUENOS AIRES 2021

We met at a (virtual) ogcapi.ogc.org sprint

GEONOVUM TESTBED
SPATIAL DATA ON THE WEB




"Web Developers"
Open Data Portals
Search Engine Crawlers
Indexed Web
Linked Data Web?
Web APIs?
GeoNetwork
DCAT
schema.org
"linked data proxy"
schema.org
WGS 84
HTML & JSON-LD
Content negotiation
"Follow your nose"
"linked data proxy"
links between datasets
Bekendmakingen
SDI
CSW
ISO 19115
ISO 19139/XML
WFS
any model
any CRS
XML
rich queries
WFS
Metadata
Adressen en Gebouwen (BAG)
Agrarisch Areaal
GIS Experts and Developers



Best Practices Summary

Best Practice 1: Use globally unique persistent HTTP URIs for Spatial Things

Best Practice 2: Make your spatial data indexable by search engines

Best Practice 3: Link resources together to create the Web of data

Best Practice 4: Use spatial data encodings that match your target audience

Best Practice 5: Provide geometries on the Web in a usable way

Best Practice 6: Provide geometries at the right level of accuracy, precision, and size

Best Practice 7: Choose coordinate reference systems to suit your user's applications

Best Practice 8: State how coordinate values are encoded

Best Practice 9: Describe relative positioning

Best Practice 10: Use appropriate relation types to link Spatial Things

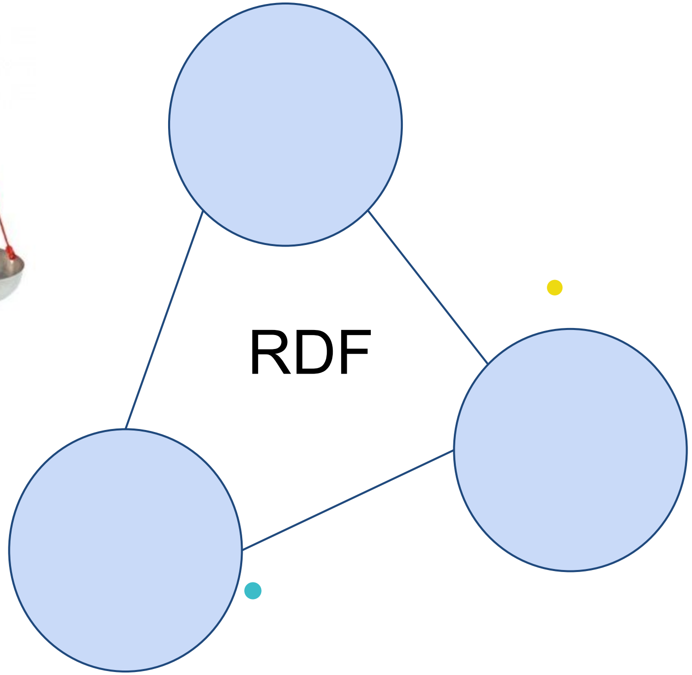
Best Practice 11: Provide information on the changing nature of spatial things

Best Practice 12: Expose spatial data through 'convenience APIs'

Best Practice 13: Include spatial metadata in dataset metadata

Best Practice 14: Describe the positional accuracy of spatial data

OWS


RDF

OGC API

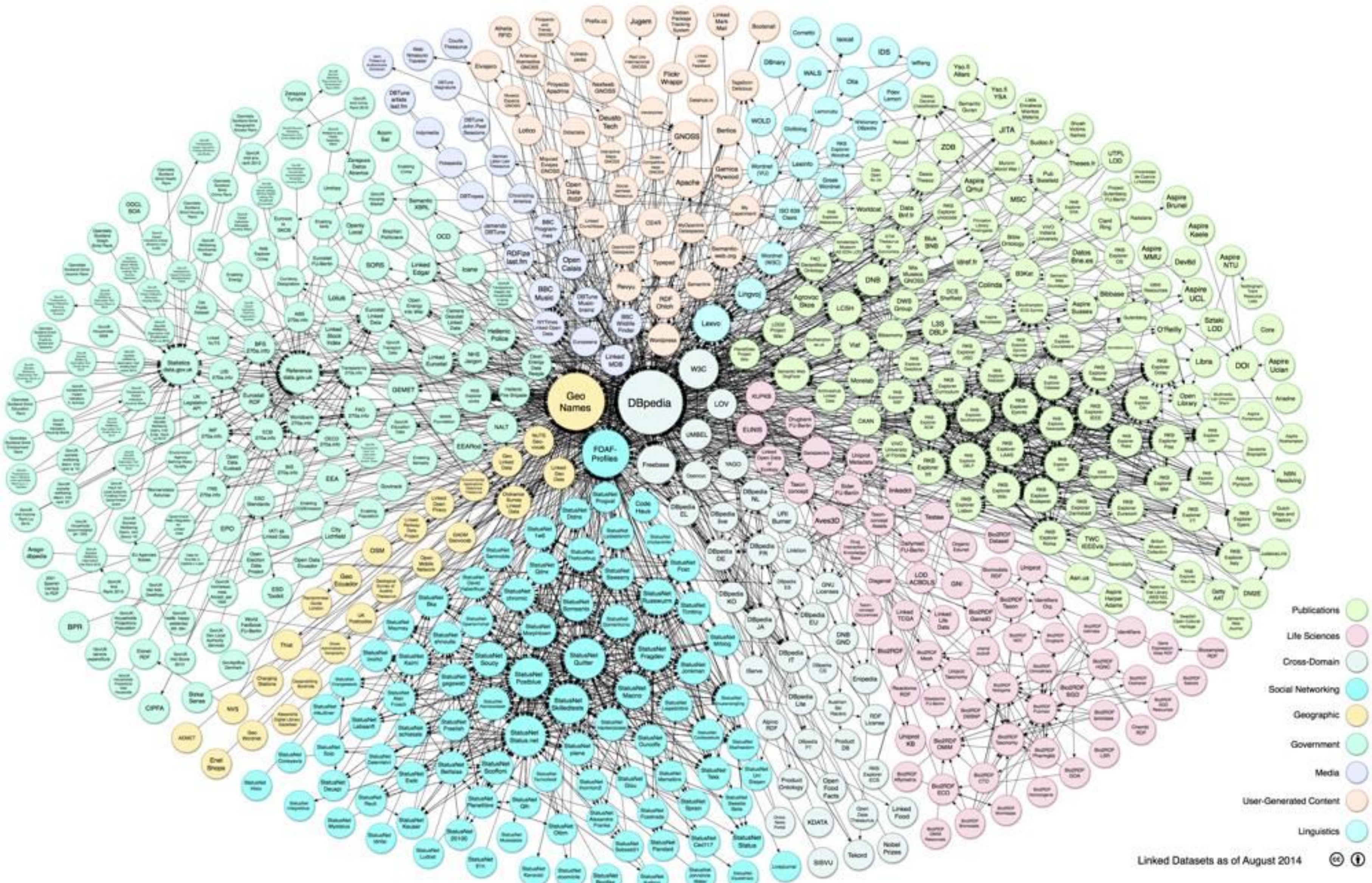
FOSS4G
BUENOS AIRES 2021

Should we add more Linked Data?

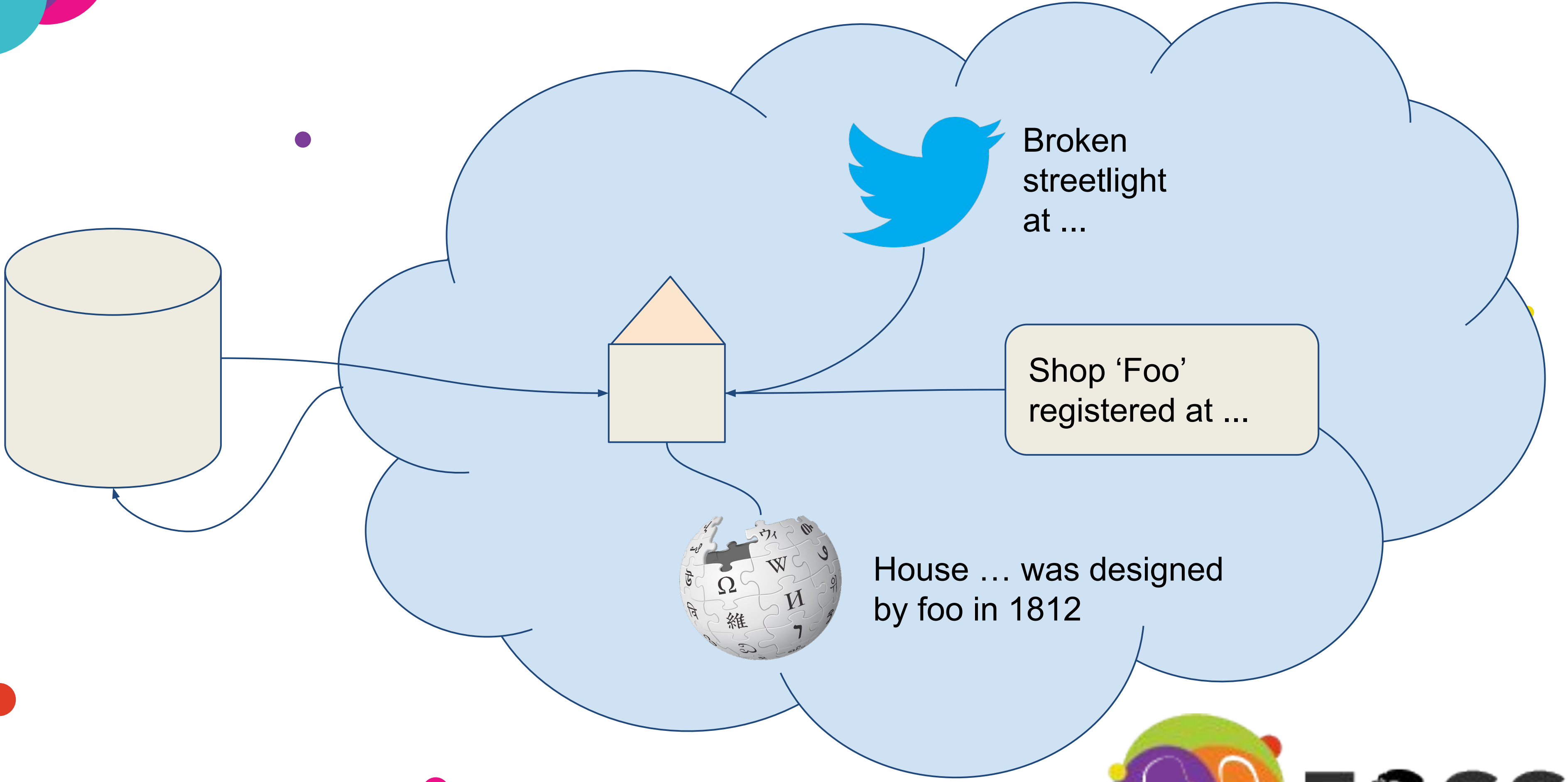
- A json-ld profile for OGC API Features / Records?
- A TTL output encoding for pygeoapi?
- A SPARQL endpoint on GeoServer?
- Let’s ask Marco!
- But first, some basic concepts of RDF

Linked data 101

Quick overview of linked data



Linked Datasets as of August 2014



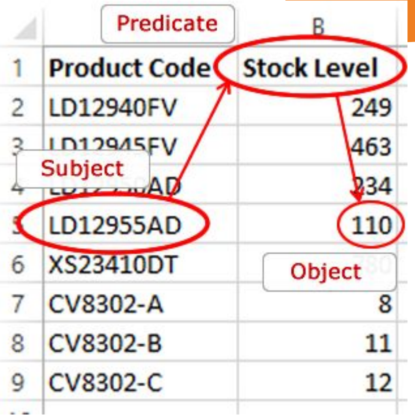
Broken streetlight at ...

Shop ‘Foo’ registered at ...

House … was designed by foo in 1812

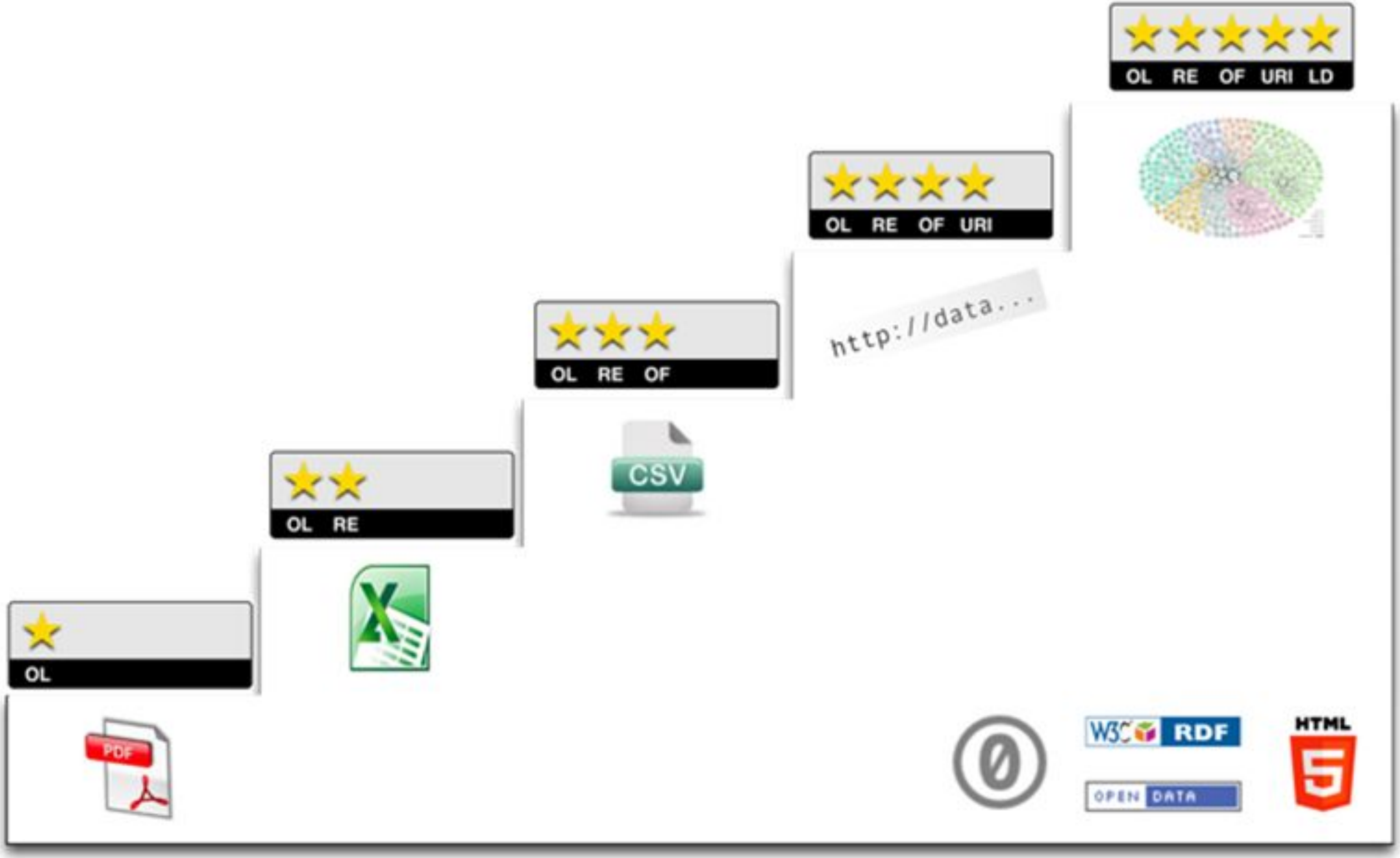
FOSS4G
BUENOS AIRES 2021

Concept	Description	In GeoSpatial
Triple	Subject Predicate Object	Feature attribute
RDF	Resource Description Framework	Relational databases
Graph	Collection of triples	Data
TripleStore	Persistence for triples	Database
URI	Identifier	ID attribute
SPARQL	Query language	SQL
Ontology / Vocabulary	Definition of a data model	Schema
OWL	Syntax to describe ontologies (Web Ontology Language)	UML / XSD
Encoding / Serialisation	RDF can be encoded as turtle, json-ld, rdf-xml, ... without loss of information	Rendering?



		B
1	Product Code	Stock Level
2	LD12940FV	249
3	LD12945FV	463
4	LD12950AD	234
5	LD12955AD	110
6	XS23410DT	[illegible]
7	CV8302-A	8
8	CV8302-B	11
9	CV8302-C	12

FOSS4G
BUENOS AIRES 2021


★
OL
PDF
★★
OL RE
★★★
OL RE OF
CSV
★★★★
OL RE OF URI
http://data...
★★★★★
OL RE OF URI LD
W3C RDF
OPEN DATA
HTML
5

4 Star; Use (W3C) standards

For example: URI’s to identify things.
https://www.w3.org/TR/cooluris

1. Be on the web
2. Be unambiguous
3. Simple, stable, manageable

- Prevent organisational/product/project names
- A webpage about a object is not the object itself

5 star; Link your data

- Use common ontologies to describe data
- Re-Use existing (wikipedia) identifiers

Use Cases

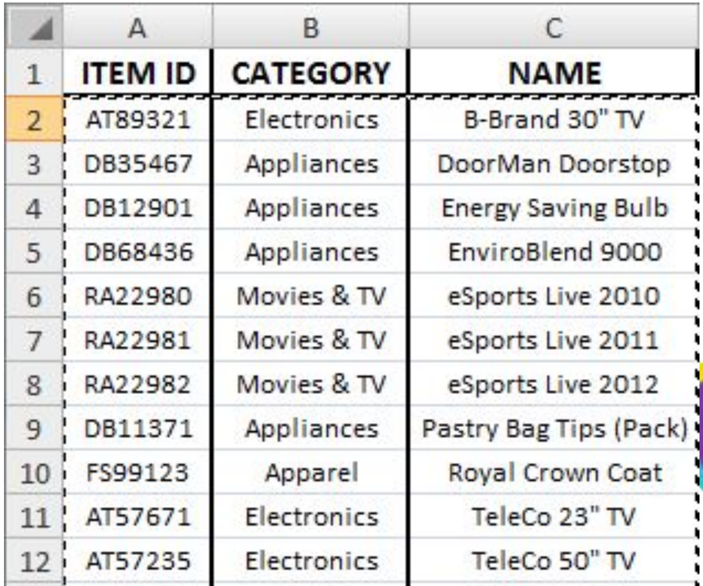
Cases for linked data in Geospatial

Case 1: Create RDF from table data?

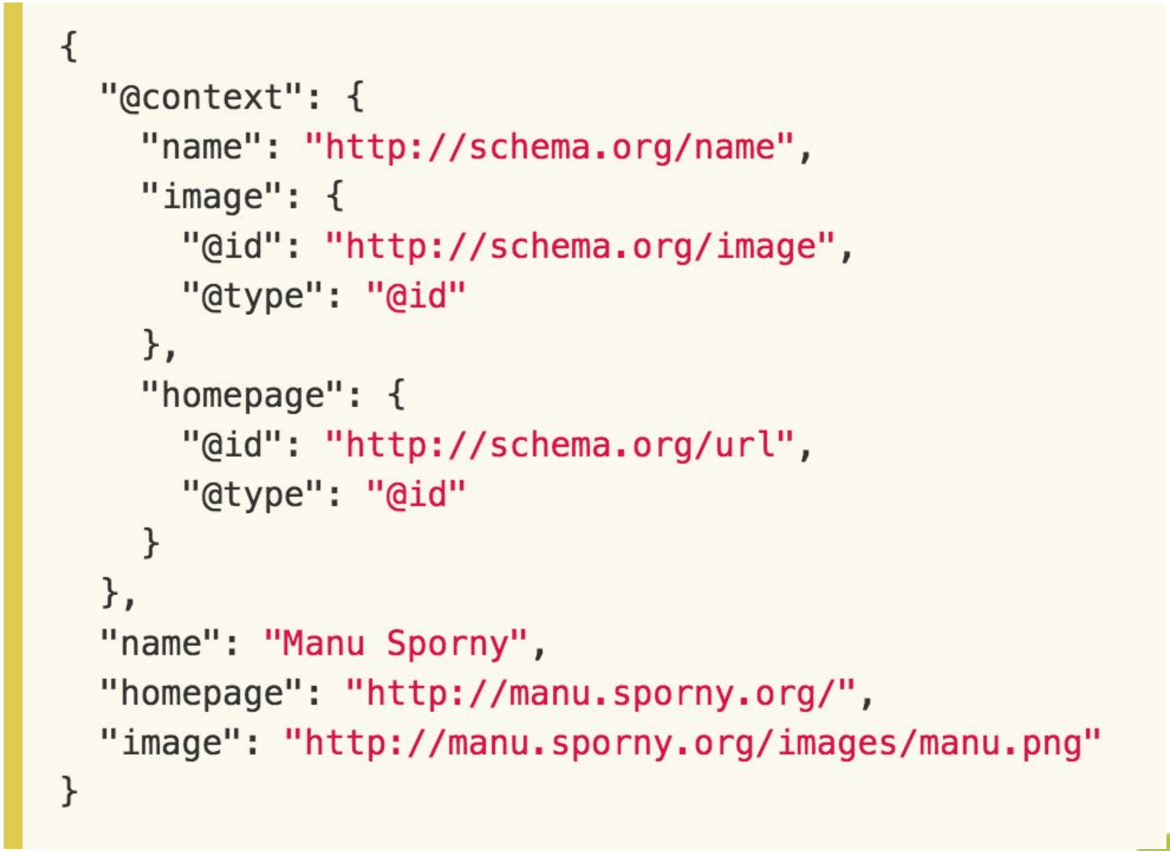
Mint a URI based on unique column schema.org/identifier

Table is type schema.org/Product

Columns as predicates from common ontologies schema.org/name



	A	B	C
1	ITEM ID	CATEGORY	NAME
2	AT89321	Electronics	B-Brand 30" TV
3	DB35467	Appliances	DoorMan Doorstop
4	DB12901	Appliances	Energy Saving Bulb
5	DB68436	Appliances	EnviroBlend 9000
6	RA22980	Movies & TV	eSports Live 2010
7	RA22981	Movies & TV	eSports Live 2011
8	RA22982	Movies & TV	eSports Live 2012
9	DB11371	Appliances	Pastry Bag Tips (Pack)
10	FS99123	Apparel	Royal Crown Coat
11	AT57671	Electronics	TeleCo 23" TV
12	AT57235	Electronics	TeleCo 50" TV



```
{
  "@context": {
    "name": "http://schema.org/name",
    "image": {
      "@id": "http://schema.org/image",
      "@type": "@id"
    },
    "homepage": {
      "@id": "http://schema.org/url",
      "@type": "@id"
    }
  },
  "name": "Manu Sporny",
  "homepage": "http://manu.sporny.org/",
  "image": "http://manu.sporny.org/images/manu.png"
}
```

Software using this approach

[Fiona](#)

A python wrapper for GDAL, has an option to convert spatial files to GeoJson-ld

[D2RQ](#)

An application that exposes a relational database as SPARQL endpoint

[pygeoapi](#)

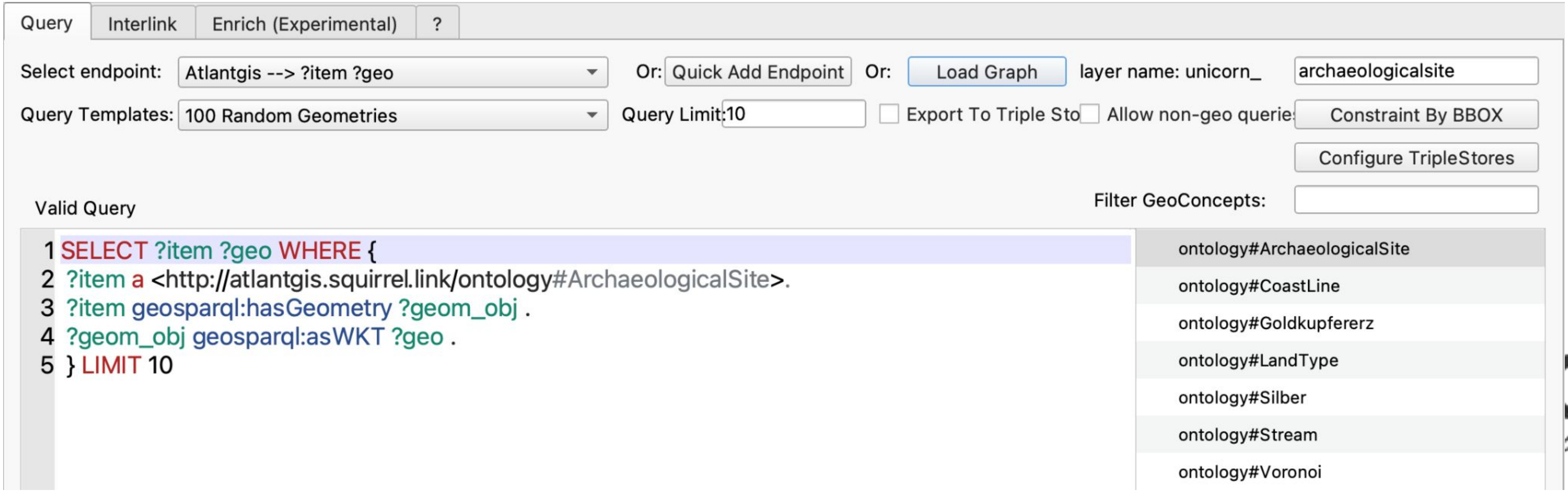
Implementation of OGC API with option to provide a ld-context for json data

[GeoServer](#)

Implementation of OGC API with option to provide a ld-context for geojson

Case 2: Consume RDF with GIS Desktop tooling

A SPARQL query as a source for a QGIS layer
https://github.com/sparqlunicorn/sparqlunicornGoesGIS



Select endpoint: Atlantgis --> ?item ?geo

layer name: unicorn_ archaeologicalsite

Query Templates: 100 Random Geometries

Query Limit:10

□ Export To Triple Sto

□ Allow non-geo querie

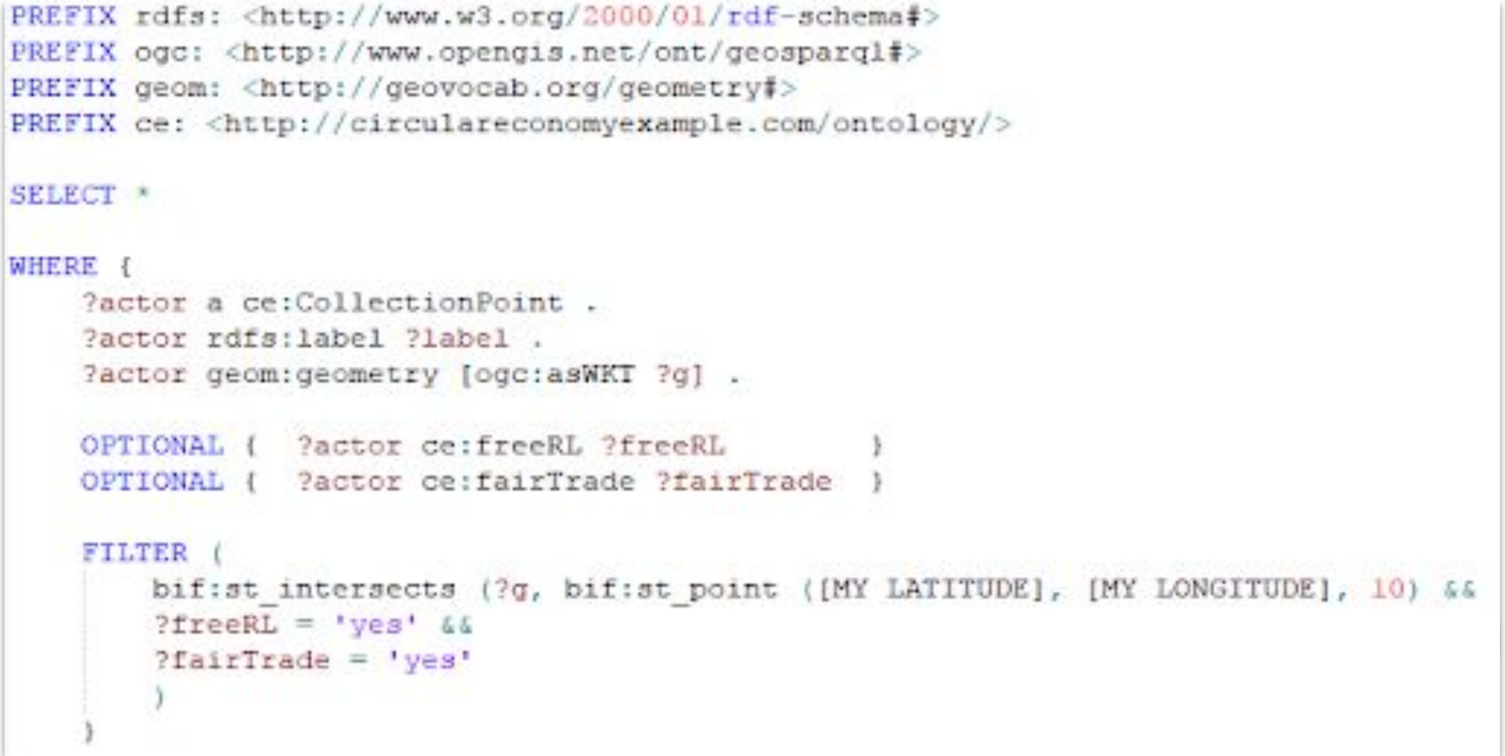
Valid Query

```
SELECT ?item ?geo WHERE {
 ?item a <http://atlantgis.squirrel.link/ontology#ArchaeologicalSite>.
 ?item geosparql:hasGeometry ?geom_obj .
 ?geom_obj geosparql:asWKT ?geo .
 } LIMIT 10
```

ontology#ArchaeologicalSite
ontology#CoastLine
ontology#Goldkupfererz
ontology#LandType
ontology#Silber
ontology#Stream
ontology#Voronoi

Case 3; Spatial analyses within the RDF context

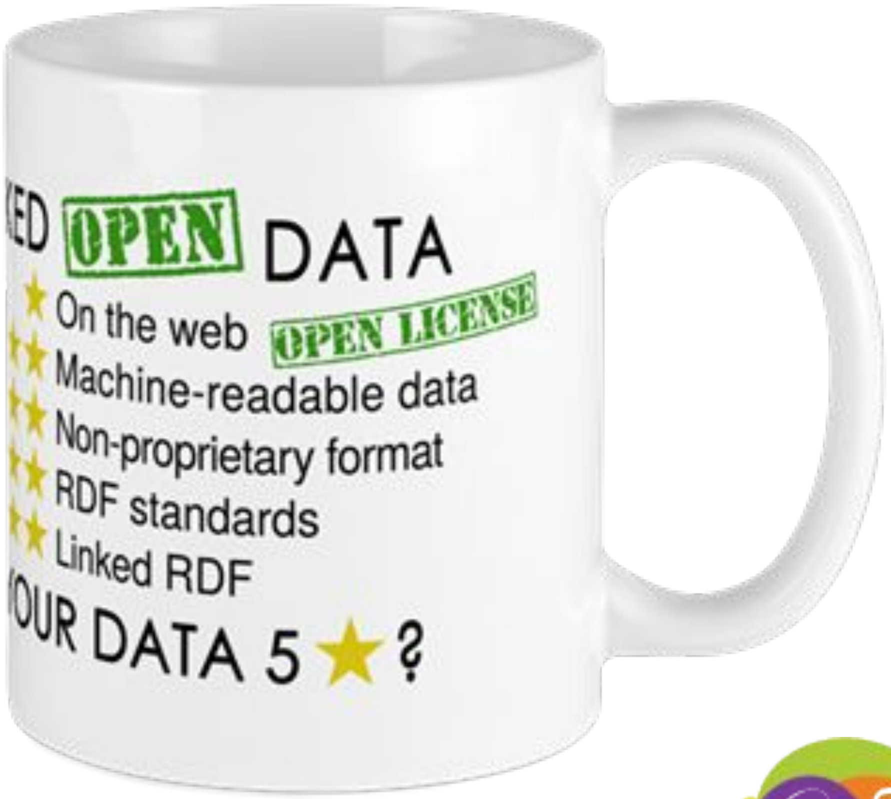
- SPARQL extended with spatial predicates and filters



```
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX ogc: <http://www.opengis.net/ont/geosparql#>
PREFIX geom: <http://geovocab.org/geometry#>
PREFIX ce: <http://circulareconomyexample.com/ontology/>

SELECT *

WHERE {
    ?actor a ce:CollectionPoint .
    ?actor rdfs:label ?label .
    ?actor geom:geometry [ogc:asWKT ?g] .

    OPTIONAL {  ?actor ce:freeRL ?freeRL          }
    OPTIONAL {  ?actor ce:fairTrade ?fairTrade    }

    FILTER (
        bif:st_intersects (?g, bif:st_point ([MY LATITUDE], [MY LONGITUDE], 10) &&
        ?freeRL = 'yes' &&
        ?fairTrade = 'yes'
        )
    }
```

GeoSPARQL

An ontology to describe (and query) spatial


KED OPEN DATA
★ On the web
OPEN LICENSE
★★ Machine-readable data
★★★ Non-proprietary format
★★★★ RDF standards
★★★★★ Linked RDF
OUR DATA 5 ★ ?

History Linked Spatial Data on The Web

SQL > Relational

SQL Spatial (Simple Features) 1999

RDF > Graph 1999

RDF Spatial 2003

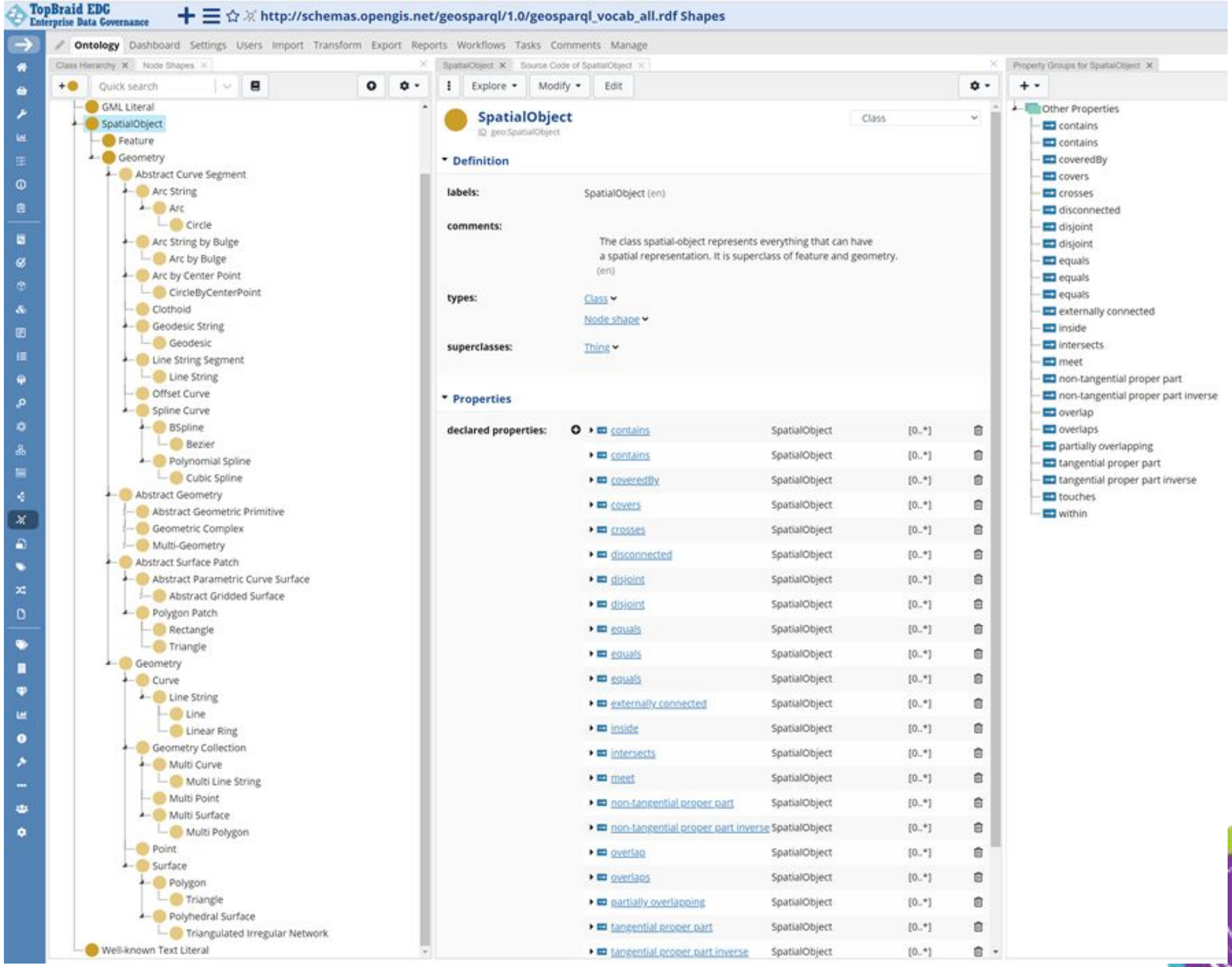
SPARQL > Spatial SPARQL Query (Jena) 2007

SPARQL > GeoSPARQL 10/2009 - 06/2012

(Full) OGC GeoSPARQL support (Jena) 06/2019

OGC GeoSPARQL

- Standard released by the OGC in 2012
- 6 Components
  - Core
  - Topology Vocabulary Extension
  - Geometry Extension
  - Geometry Topology Extension



- GML Literal
- SpatialObject
  - Feature
  - Geometry
    - Abstract Curve Segment
      - Arc String
        - Arc
          - Circle
      - Arc String by Bulge
        - Arc by Bulge
      - Arc by Center Point
        - CircleByCenterPoint
      - Clothoid
      - Geodesic String
        - Geodesic
      - Line String Segment
        - Line String
      - Offset Curve
      - Spline Curve
        - BSpline
          - Bezier
        - Polynomial Spline
          - Cubic Spline
    - Abstract Geometry
      - Abstract Geometric Primitive
      - Geometric Complex
      - Multi-Geometry
    - Abstract Surface Patch
      - Abstract Parametric Curve Surface
        - Abstract Gridded Surface
      - Polygon Patch
        - Rectangle
        - Triangle
    - Geometry
      - Curve
        - Line String
          - Line
          - Linear Ring
      - Geometry Collection
        - Multi Curve
          - Multi Line String
        - Multi Point
        - Multi Surface
          - Multi Polygon
      - Point
      - Surface
        - Polygon
          - Triangle
        - Polyhedral Surface
          - Triangulated Irregular Network
- Well-known Text Literal

SpatialObject

ID geo:SpatialObject

Definition

labels: SpatialObject (en)

comments: The class spatial-object represents everything that can have a spatial representation. It is superclass of feature and geometry. (en)

types: Class
Node shape

superclasses: Thing

Properties

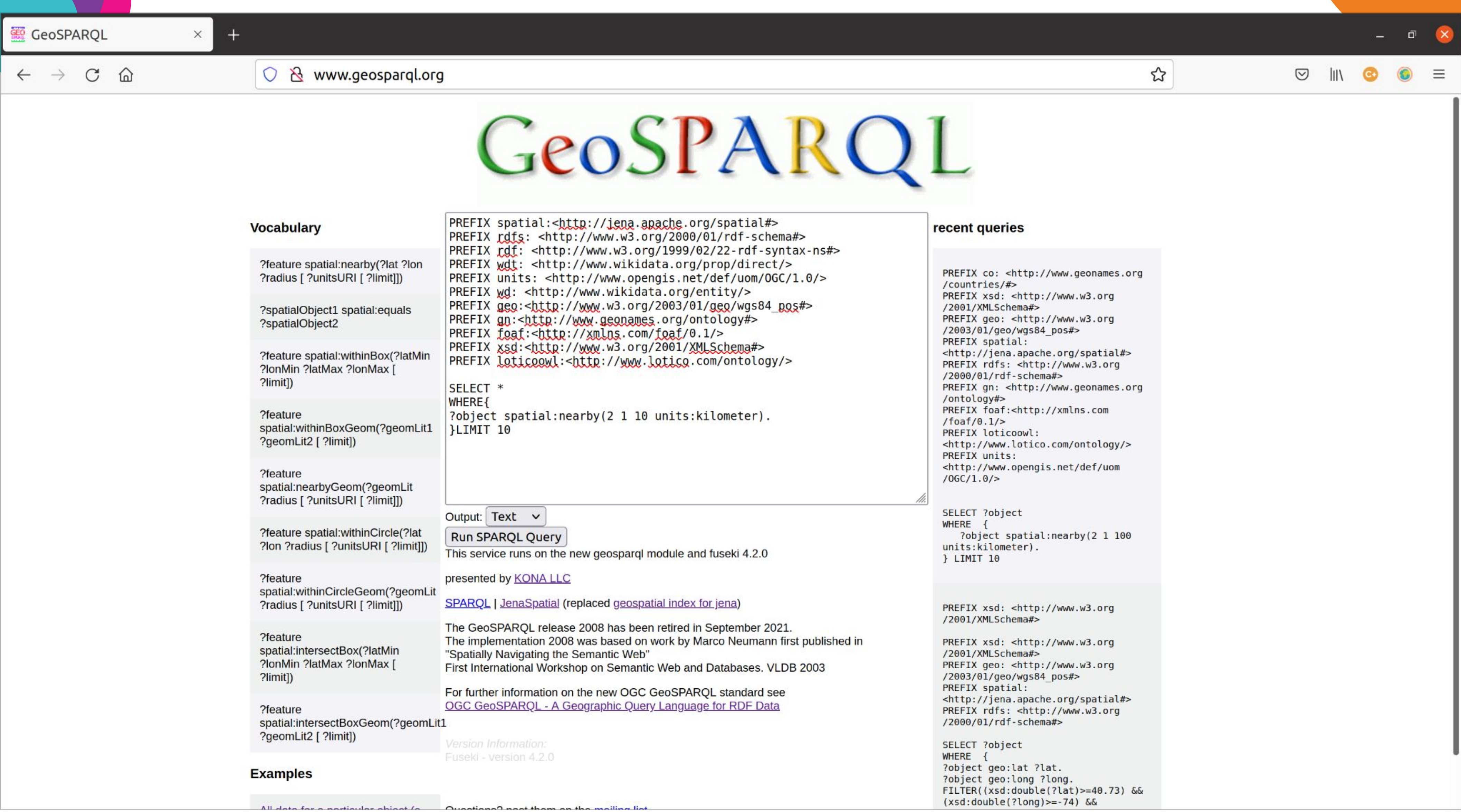
declared properties:

contains	SpatialObject	[0..*]
contains	SpatialObject	[0..*]
coveredBy	SpatialObject	[0..*]
covers	SpatialObject	[0..*]
crosses	SpatialObject	[0..*]
disconnected	SpatialObject	[0..*]
disjoint	SpatialObject	[0..*]
disjoint	SpatialObject	[0..*]
equals	SpatialObject	[0..*]
equals	SpatialObject	[0..*]
equals	SpatialObject	[0..*]
externally connected	SpatialObject	[0..*]
inside	SpatialObject	[0..*]
intersects	SpatialObject	[0..*]
meet	SpatialObject	[0..*]
non-tangential proper part	SpatialObject	[0..*]
non-tangential proper part inverse	SpatialObject	[0..*]
overlap	SpatialObject	[0..*]
overlaps	SpatialObject	[0..*]
partially overlapping	SpatialObject	[0..*]
tangential proper part	SpatialObject	[0..*]
tangential proper part inverse	SpatialObject	[0..*]

- Other Properties
  - contains
  - contains
  - coveredBy
  - covers
  - crosses
  - disconnected
  - disjoint
  - disjoint
  - equals
  - equals
  - equals
  - externally connected
  - inside
  - intersects
  - meet
  - non-tangential proper part
  - non-tangential proper part inverse
  - overlap
  - overlaps
  - partially overlapping
  - tangential proper part
  - tangential proper part inverse
  - touches
  - within

OGC GeoSPARQL

- Standard released by the OGC in 2012
- 6 Components
  - Core
  - Topology Vocabulary Extension
  - Geometry Extension
  - Geometry Topology Extension
  - RDFS Entailment Extension
  - Query Rewrite Extension



Vocabulary

?feature spatial:nearby(?lat ?lon ?radius [?unitsURI [?limit]])

?spatialObject1 spatial:equals ?spatialObject2

?feature spatial:withinBox(?latMin ?lonMin ?latMax ?lonMax [?limit])

?feature spatial:withinBoxGeom(?geomLit1 ?geomLit2 [?limit])

?feature spatial:nearbyGeom(?geomLit ?radius [?unitsURI [?limit]])

?feature spatial:withinCircle(?lat ?lon ?radius [?unitsURI [?limit]])

?feature spatial:withinCircleGeom(?geomLit ?radius [?unitsURI [?limit]])

?feature spatial:intersectBox(?latMin ?lonMin ?latMax ?lonMax [?limit])

?feature spatial:intersectBoxGeom(?geomLit1 ?geomLit2 [?limit])

Examples

All data for a particular object (a

```
PREFIX spatial:<http://jena.apache.org/spatial#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX wdt: <http://www.wikidata.org/prop/direct/>
PREFIX units: <http://www.opengis.net/def/uom/OGC/1.0/>
PREFIX wd: <http://www.wikidata.org/entity/>
PREFIX geo:<http://www.w3.org/2003/01/geo/wgs84_pos#>
PREFIX gn:<http://www.geonames.org/ontology#>
PREFIX foaf:<http://xmlns.com/foaf/0.1/>
PREFIX xsd:<http://www.w3.org/2001/XMLSchema#>
PREFIX loticoowl:<http://www.lotico.com/ontology/>

SELECT *
WHERE{
?object spatial:nearby(2 1 10 units:kilometer).
}LIMIT 10
```

Output: Text

Run SPARQL Query

This service runs on the new geosparql module and fuseki 4.2.0

presented by KONA LLC

SPARQL | JenaSpatial (replaced geospatial index for jena)

The GeoSPARQL release 2008 has been retired in September 2021.
The implementation 2008 was based on work by Marco Neumann first published in "Spatially Navigating the Semantic Web"
First International Workshop on Semantic Web and Databases. VLDB 2003

For further information on the new OGC GeoSPARQL standard see
OGC GeoSPARQL - A Geographic Query Language for RDF Data

Version Information:
Fuseki - version 4.2.0

Questions? post them on the mailing list

recent queries

```
PREFIX co: <http://www.geonames.org
/countries/#>
PREFIX xsd: <http://www.w3.org
/2001/XMLSchema#>
PREFIX geo: <http://www.w3.org
/2003/01/geo/wgs84_pos#>
PREFIX spatial:
<http://jena.apache.org/spatial#>
PREFIX rdfs: <http://www.w3.org
/2000/01/rdf-schema#>
PREFIX gn: <http://www.geonames.org
/ontology#>
PREFIX foaf:<http://xmlns.com
/foaf/0.1/>
PREFIX loticoowl:
<http://www.lotico.com/ontology/>
PREFIX units:
<http://www.opengis.net/def/uom
/OGC/1.0/>


SELECT ?object
WHERE  {
   ?object spatial:nearby(2 1 100
units:kilometer).
} LIMIT 10
```

```
PREFIX xsd: <http://www.w3.org
/2001/XMLSchema#>

PREFIX xsd: <http://www.w3.org
/2001/XMLSchema#>
PREFIX geo: <http://www.w3.org
/2003/01/geo/wgs84_pos#>
PREFIX spatial:
<http://jena.apache.org/spatial#>
PREFIX rdfs: <http://www.w3.org
/2000/01/rdf-schema#>

SELECT ?object
WHERE  {
?object geo:lat ?lat.
?object geo:long ?long.
FILTER((xsd:double(?lat)>=40.73) &&
(xsd:double(?long)>=-74) &&
```

Q & A

Thank you for your attention