



**HAL**  
open science

## ‘Linked data’ the glue within interoperable information systems

Sylvain Grellet, Abdelfettah Feliachi

► **To cite this version:**

Sylvain Grellet, Abdelfettah Feliachi. ‘Linked data’ the glue within interoperable information systems. INSPIRE Conference 2020, Jun 2020, Dubrovnik, Croatia. hal-03499535

**HAL Id: hal-03499535**

**<https://brgm.hal.science/hal-03499535>**

Submitted on 21 Dec 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



# 'LINKED DATA' THE GLUE WITHIN INTEROPERABLE INFORMATION SYSTEMS

---

ABDEFETTAH FELIACHI  
SYLVAIN GRELLET



Geoscience for a sustainable Earth

**brgm**

# CONTEXT

‘Previous generation’ of OGC APIs (ex: WFS, SOS)

based on design from the 90's, 2000 (RPC, SOAP)

## Brought SDIs to the web but

- Are opaque to the web (and crawlers)
  - need to know the catalogue where to search for the API end-point
  - need to know how to interact with the such catalogue
  - when the API end-point is found, need to RTFM of the API spec

=> We expose(d) data in a siloed way

Idea n°1: Data is naturally linked  
-> why not linking data content ?

# (1) ? LINKING DATA CONTENT ?

## Still using 'Previous generation' OGC APIs

- Assigning HTTP URIs
  - Applying W3C best practices (Data on the web, Spatial Data on the web)
  
- Links
  - Between features
    - Station to river,
    - AquiferSystem to its part, ...
  - To controlled vocabularies

# (1) ? LINKING DATA CONTENT ?

- Example

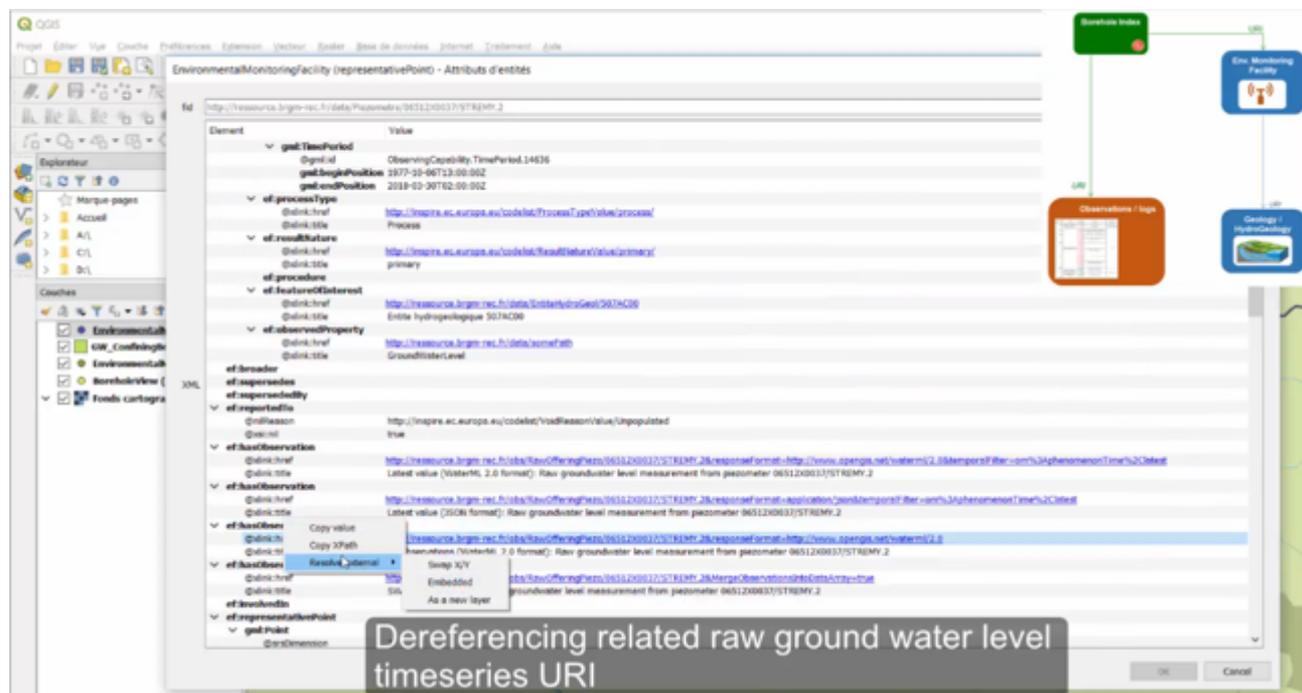
```
<gwfml2:GW_AquiferSystem gml:id="EntiteHydroGeol.121AS">
  <gml:description>Aquifer from national BD LISA (Base de Données des Limites des Systèmes Aquifères) database</gml:description>
  <gml:identifier codeSpace="http://www.ietf.org/rfc/rfc2616">https://data.geoscience.fr/id/hydrogeounit/121AS</gml:identifier>
  <gml:name codeSpace="http://id.eaufrance.fr/isaq/2/"> Craie du Sénonien au Turonien inférieur du Bassin Parisien du Neubourg au Saint-André - bassin versant de l'Eure, Iton, Avre (ba
  <gsmlb:observationMethod>
    <swe:Category>
  </gsmlb:observationMethod>
  <gsmlb:occurrence>
  <gsmlb:purpose>instance</gsmlb:purpose>
  <gsmlb:geologicUnitType xlink:href="http://resource.geosciml.org/classifier/cgi/geologicunittype/hydrogeologic_unit" xlink:title="Hydrogeologic Unit"/>
  <gsmlb:rank xlink:href="http://id.eaufrance.fr/nsa/352#2" xlink:title="Régional"/>
  <gwfml2:gwUnitMedia xlink:href="http://id.eaufrance.fr/nsa/353#9" xlink:title="Triple porosité : matricielle et karstique et de fractures"/>
  <gwfml2:gwAquiferSystem xlink:href="https://data.geoscience.fr/id/hydrogeounit/121" xlink:title="Grand système multicouche du Campanien au Turonien (Seno-Turonien) du Bassin Parisien du Neubourg au Saint-André - bassin versant de l'Eure, Iton, Avre (bassin Seine-Normandie)"/>
  <gwfml2:gwAquiferSystemPart xlink:href="https://data.geoscience.fr/id/hydrogeounit/121AS01" xlink:title="Craie du Séno-Turonien du Bassin Parisien du Neubourg au Saint-André - bassin versant de l'Eure, Iton, Avre (bassin Seine-Normandie)"/>
  <gwfml2:gwAquiferSystemPart xlink:href="https://data.geoscience.fr/id/hydrogeounit/121AS30" xlink:title="Craie marseuse et marnes du Turonien inférieur du Bassin Parisien du Neubourg au Saint-André - bassin versant de l'Eure, Iton, Avre (bassin Seine-Normandie)"/>
</gwfml2:GW_AquiferSystem>
```

URI to Feature

URI to codeList

# (1) ? LINKING DATA CONTENT ?

- From this you can already build nice things



See :

[https://github.com/BRGM/gml\\_application\\_schema\\_toolbox/tree/master/presentations/2018\\_INSPIRE\\_conference](https://github.com/BRGM/gml_application_schema_toolbox/tree/master/presentations/2018_INSPIRE_conference)

Even between 'Previous generation' OGC APIs that's already a better experience

# (1) ? LINKING DATA CONTENT ?

## Still using 'Previous generation' OGC APIs

- Data content starts to open up
  - -> start to cross silos
  - -> that's already a better experience
- But
  - API are still opaque to the web
  - Data semantics unaccessible to the web of data

Idea n°2: Why not serving data in a web oriented approach ?

## (2) ? SERVING DATA IN A WEB ORIENTED APPROACH ?

### New OGC APIs

- Important turn/change in OGC specs
  - Web APIs, sometimes referred to as RESTful services.
  - Each API instance described by a documentation that is meaningful for the Web
  - No retrocompatibility is not an issue
  - Targetting « JSON family » (often (Geo)JSON) representation
  
- Hypermedia (HATEOAS)
  - Within API operation
  - Within data structure returned: perfect example of ST API response
  - -> A client could work by following links between resources, that's the way of the modern Web



## (2) ? SERVING DATA IN A WEB ORIENTED APPROACH ?

### New OGC APIs

- SensorThings API

```
{
  "@iot.nextLink": "https://sensorthings-wq.brgm-rec.fr/FROST-Server/v1.0/Datastreams?$skip=100",
  "value": [
    {
      "name": "Dichloropropène-2,3 in Eau on Eau brute at LA GRANDE NOUE D'HERMÉ A GOUAIX 1 with method Qualité de l'eau - Dosage par chromatographie en phase gazeuse d'un certain nombre d'hydrocarbures aromatiques monocycliques, du naphthalène et de divers composés chlorés par dégazage, piégeage et désorption thermique (T90-129/NF EN ISO 15680 01/2004)",
      "description": "Dichloropropène-2,3 in Eau on Eau brute at LA GRANDE NOUE D'HERMÉ A GOUAIX 1 with method Qualité de l'eau - Dosage par chromatographie en phase gazeuse d'un certain nombre d'hydrocarbures aromatiques monocycliques, du naphthalène et de divers composés chlorés par dégazage, piégeage et désorption thermique (T90-129/NF EN ISO 15680 01/2004)",
      "properties": {↔},
      "observationType": "http://www.opengis.net/def/observationType/OGC-OM/2.0/OM_Measurement",
      "observedArea": {
        "type": "Point",
        "coordinates": [
          48.457967041,
          3.296382646
        ]
      }
    },
    {
      "phenomenonTime": "2010-01-04T00:20:00.000Z/2011-11-02T11:55:00.000Z",
      "Sensor@iot.navigationLink": "https://sensorthings-wq.brgm-rec.fr/FROST-Server/v1.0/Datastreams(1142553)/Sensor",
      "Thing@iot.navigationLink": "https://sensorthings-wq.brgm-rec.fr/FROST-Server/v1.0/Datastreams(1142553)/Thing",
      "Observations@iot.navigationLink": "https://sensorthings-wq.brgm-rec.fr/FROST-Server/v1.0/Datastreams(1142553)/Observations",
      "unitOfMeasurement": {↔},
      "ObservedProperty@iot.navigationLink": "https://sensorthings-wq.brgm-rec.fr/FROST-Server/v1.0/Datastreams(1142553)/ObservedProperty",
      "@iot.id": 1142553,
      "@iot.selfLink": "https://sensorthings-wq.brgm-rec.fr/FROST-Server/v1.0/Datastreams(1142553)"
    }
  ]
}
```

## (2) ? SERVING DATA IN A WEB ORIENTED APPROACH ?

### New OGC APIs

- APIs endpoint can be automatically discovered (by machines)
- But
  - Data semantics unaccessible to the web of data

Idea n°3: Why not expressing data semantics differently ?

## (3) ? EXPRESSING DATA SEMANTICS DIFFERENTLY ?

### Using web of data oriented semantics

- OGC ELFIE / SELFIE (<https://opengeospatial.github.io/ELFIE/> )
  - 2018 – 2019 ELFIE: OGC and W3C standards with the goal of establishing a best practice for exposing links between and among environmental domain and sampling features in a highly adaptable standards compliant way that is compatible with modern web search technology.
  - 2019 – 2020 SELFIE: designing and vetting Web-resource model and network behavior for cross-domain linked feature data that compliments and uses new OGC APIs as a building block.
- JSON-LD serialization/model of choice
- Semantics
  - Schema.org for indexing by search engines
  - Complemented by ontologies stemming from OGC, INSPIRE semantics, W3C, ...
- Content negotiation on HTTP URIs

# (3) ? EXPRESSING DATA SEMANTICS DIFFERENTLY ?

## Using web of data oriented semantics

- Controlled vocabulary Registry
  - Being federated at the EU level

The image shows a screenshot of the BRGM Geoscience Registry website. The top part displays the 'Outil de gestion des registres du BRGM' with navigation links and a search bar. Below this, the 'Registre: Lithologie' is shown, including its URI and a brief description. A metadata table is visible on the right side of the page.

The bottom part of the image shows the 'Register Federation' interface, which lists various registries. The 'Registries' table is as follows:

Filter Label	Filter Publisher	Filter URI
<b>Label</b>	<b>Publisher</b>	<b>URI</b>
European Geoscience Registry	BRGM	https://data.geoscience.earth/nci/
Système de publication de registres de l'IDG française	BRGM	http://registre.geocatalogue.fr/registry
INSPIRE registry	European Commission, Joint Research Centre	http://inspire.ec.europa.eu/registry
EEA Data Dictionary vocabularies	European Environment Agency	http://dd.elonet.europa.eu/vocabularies
Austrian INSPIRE registry	CCCA Data Centre	http://registry.inspire.gov.at
INSPIRE Italia Registry	Agency for Digital Italy	http://registry.geodati.gov.it/registry
Lithological Classification Austria - extended values	Geological Survey of Austria	http://resource.geolba.ac.at/lithology/0
Spanish INSPIRE registry	Centro Nacional de Información Geográfica	http://registro.idee.es/registry
BRGM (French geological survey) registry tool	BRGM	https://data.geoscience.fr/nci/
IUGS CGI Geoscience Vocabularies for Linked Data	Geoscience Terminology Working Group (GTWG)	http://resource.geosciml.org/

At the bottom right, the BRGM logo is displayed with the tagline 'Geoscience for a sustainable Earth'.

# (3) ? EXPRESSING DATA SEMANTICS DIFFERENTLY ?

## Using web of data oriented semantics

- Feature API support



[About](#) | [Blog](#) | [Download](#)

GeoServer 2.18.x User Manual » Community modules » JSON-LD Extension

[previous](#) | [next](#) | [modules](#)

## JSON-LD Extension

This plugin brings in the ability to output features in a json-ld format, using a json-ld file as a template to operate a second level mapping over such features. [JSON-LD](#) is a Linked Data format, based on JSON format, and revolves around the concept of "context" to provide additional mappings from JSON to an RDF model.

JSON-LD extension works for both simple and complex-features.

- [Installing the GeoServer JSON-LD extension](#)
- [JSON-LD Configuration](#)
- [JSON-LD Output Format](#)
- [JSON-LD Backward Mapping](#)

geopython / pygeoapi

Watch 20 Star 106 Fork 59

<> Code Issues 79 Pull requests 11 Actions Projects 3 Wiki Security 0 Insights

## Add JSON-LD representations #246

Open with

Merged tomkralidis merged 37 commits into geopython:master from spatialdataotearoa:ld-json on 19 Dec 2019

Conversation 36 Commits 37 Checks 0 Files changed 12 +554 -118



alpha-beta-soup commented on 24 Sep 2019

Contributor

This is something I've been considering for a little while. It's great for SEO that pygeoapi includes microdata read from the config in the HTML templates. However I think we could do one better by supporting JSON-LD representations of the API itself, and also extensible GeoJSON-LD representations of actual data.

Reviewers

- tomkralidis
- pvgenuchten
- jorgejesus

## (3) ? EXPRESSING DATA SEMANTICS DIFFERENTLY ?

### Using web of data oriented semantics

- Catalogue API support

#### ISO19139 / Formatter / JSON-LD #3714

Merged fxprunayre merged 7 commits into geonetwork:master from fxprunayre:feature/Iso19139/formatter/jsonld on 11 Jun 2019

Conversation 4 Commits 7 Checks 0 Files changed 6



fxprunayre commented on 29 Mar 2019

Member

Convert an ISO19139 records in JSON-LD format.  
This JSON-LD can be embedded in an HTML page using

```
<html>
  <script type="application/ld+json">
    {json-ld}
  </script>
</html>
```

Based on <https://schema.org/Dataset> and mapping proposed by BRGM team  
<https://github.com/geonetwork/core-geonetwork/wiki/JSON-LD---ISO19139-mapping-proposal>.

Reviewers

pvgenuchten

Assignees

No one assigned

Labels

enhancement

Projects

None yet

- SensorThings API
  - Discussed in SELFIE

### (3) ? EXPRESSING DATA SEMANTICS DIFFERENTLY ?

#### Using web of data oriented semantics

- Linked data client

BLIV BRGM Linked Data Viewer

Viewer Examples About

https://opengeospatial.github.io/ELRE/FR/Borehole/gwi/B55000 Launch Add

URI dereferencing O SPARQL Endpoint  
Specify the expected data format

JSON-LD Turtle RDF/XML

```
{
  "@sampleOf": "https://sweet.jpl.nasa.gov/2.3/iri/EarthReference.owl#EarthGeosphere",
  "@featureOfInterestOf": [
    {
      "@context": "https://opengeospatial.github.io/ELRE/json-ld/foosa.jsonld",
      "@id": "http://resource.brgm-rec.fr/obs/RawGeologicLogs/B55000DXBD",
      "@type": "Observation",
      "hasFeatureOfInterest": "https://opengeospatial.github.io/ELRE/FR/Borehole/gwi/B55000DXBD",
      "hasResult": [
        {
          "@id": "http://resource.brgm-rec.fr/obs/RawGeologicLogs/B55000DXBD",
          "@type": "GW_GeologyLog"
        }
      ]
    }
  ]
}
```

From JSON-LD

https://opengeospatial.github.io/ELRE/FR/WaterBody/gwi/E6390700

Stabilize graph

Types

- Concept
- none
- Observation
- Sample
- GW\_Well
- Geometry
- Measurement
- Geometry

<https://data.geoscience.fr/Bliv/>

- Benefits

- Data instances not hidden anymore behind the API (as with previous OGC APIs)
  - > data instances can be indexed by search engines
  - > more visibility (data back on the 1<sup>st</sup> line), more reuse of public environmental data/assets

# GENERALISATION

## What a journey !

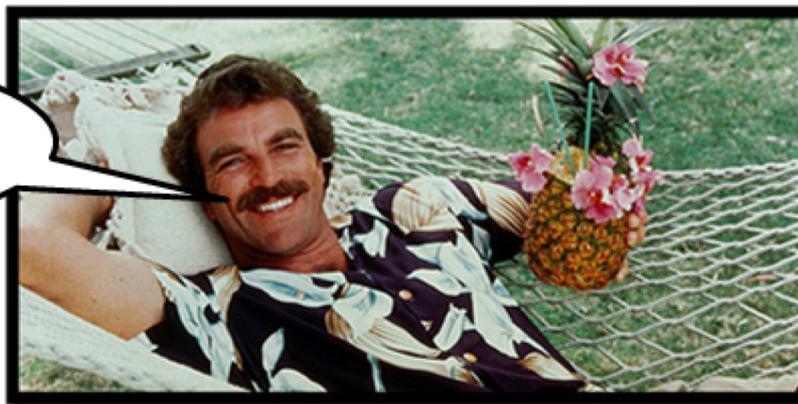
- Previous era/decade was bringing SDI's to the Web
  - INSPIRE dynamics abide by this (previous OGC APIs, INSPIRE network services guidelines)
- Now SDI's (or environmental IS) are starting to contribute to the web of data
  - Thanks to Linked Data & SemanticWeb
- Note : could have detailed "Metadata" dynamics as well OGC API Record, DCAT
  - But other presentations in that session speak about this
- OGC, INSPIRE related ontologies
  - Generated during (S)ELFIE, being published by OGC-NA
- HTTP URIs on everything
  - Features, codeList entries, APIs, dataset ... to link everything

# CONCLUSION

## Introducing LinkedData and SemanticWeb in Information Systems

- Data providers still provide their data
  - in a way that allows both APIs and data to be discovered and linked to
- > more reuse
- > less loss of time searching for, trying to understand data instead of producing new domain content
- > more domain experts doing their work

Now I have a taste  
of INSPIRE



## LINKS COMPENDIUM

- W3C Best practices
  - Data on the Web Best Practices : <https://www.w3.org/TR/dwbp/>
  - Spatial Data on the Web Best Practices : <https://www.w3.org/TR/sdw-bp/>
- QGIS GMLAS Toolbox
  - [https://github.com/BRGM/gml\\_application\\_schema\\_toolbox](https://github.com/BRGM/gml_application_schema_toolbox)
- OGC Linked data Interoperability Experiments
  - ELFIE : <https://opengeospatial.github.io/ELFIE/>
  - SELFIE: <https://github.com/opengeospatial/SELFIE>
- Registry tool - UkGovLd
  - <https://github.com/UKGovLD>
- INSPIRE register federation
  - <https://inspire-regadmin.jrc.ec.europa.eu/ror/>
- GeoServer JSON-LD community plugin
  - <https://docs.geoserver.org/latest/en/user/community/json-ld/index.html>
- PyGeoApi – JSON-LD
  - <https://github.com/geopython/pygeoapi/pull/246>
- Geonetwork – JSON-LD
  - <https://github.com/geonetwork/core-geonetwork/pull/3714>

THANK YOU

[a.feliachi@brgm.fr](mailto:a.feliachi@brgm.fr)  
[s.grellet@brgm.fr](mailto:s.grellet@brgm.fr)

