BRGM

JSON-LD Output Functionalities Extension Final Report

Nuno Oliveira
Marco Volpini
GeoSolutions SAS
15/09/2020
Version 01.01

Table of Contents

Introduction

The document *BRGM - JSON-LD Output Functionalities Extension - Technical Proposal - V.01.03* provides a detailed technical description of the enhancements that were implemented.

Handling the aquifers use case requires more control over where each feature produced by App-Schema will end up in the produced JSON-LD document, the following enhancements were implemented:

- filtering support on the JSON-LD template,
- backwards filtering support on filtered features,
- parameterization of the JSON-LD templating,
- on demand semantic validation of JSON-LD, and
- implementation of the aquifers use case.

In this document we will provide concrete examples for all the enhancements that were implemented using the aquifers use case. All example requests will target GeoSolutions BRGM DEV server.

What's need to use this functionalities

In order to use the functionality the following GeoServer plugins are needed:

- App-schema plugin
- OGCAPI plugin
- WFS-templating plugin

Target GeoServer official releases

The implemented enhancements will be available in the following upcoming GeoServer official releases, they where not backported to the GeoServer 2.16.x branch:

- 2.17.3, tentative planning for the 19 September 2020
- 2.18.0, tentative planning for the 19 September 2020

GeoServer 2.16.x branch is reaching its end of life and recent enhancements and fixes, we depend on, for the OGC API Features plugin where not backported to 2.16.x.

Implementing the aquifers use case

The aquifers use case is based on a generic domain model (see Figure 1.0) that allows us to represent relationships between features, implementing this domain model in App-Schema and the respective JSON-LD output raised a few challenges.

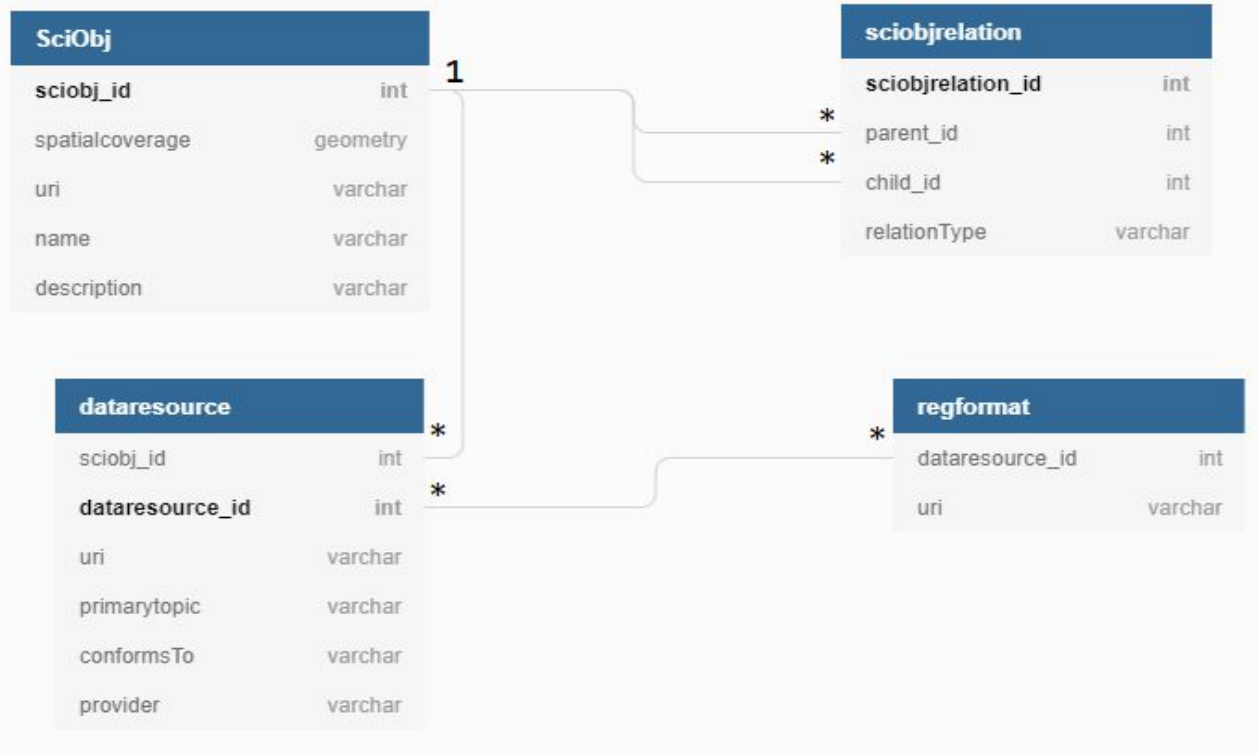

Figure 1.0 - Domain model behind the aquifers use case.

The domain model behind the aquifers use case has the following properties:

- Elements in the table *sciobj* have a many-to-many parent child relationship between them; the relationship is mapped by the table *sciobjrelation*.
- Every relation has a type value, either *dct:hasPart* or *dct:isPartOf*.
- *Sciobj* also has a one-to-many relationship with the *dataresource* table.
- Table *dataresource* has a many-to-many relationship with the *reg_format* table. To support it in app-schema a denormalized view has been created showing the *dataresource_id* and the *format* value (see SQL query 1.0).

The final JSON-LD output is expected to show for each feature the child elements in the proper JSON array, according to the *linkeddatagazetteer.sciobjrelation* type value namely the *isPartOf*

JSON array for those features having the *relationType* = *'dct:isPartOf'* (see Listing 1.0). It means that the template mechanism needs the capability to discern if it has to encode or not an element based on the value of some feature property.

```
{
 "@id":"SciObj.7290",
 "@type":[
   "Feature",
   "gw:GW_Aquifer"
 ],
 "name":{
   "@language":"fr",
   "@value":"Dépôts holocènes en Pays de la Loire et en Poitou-Charentes"
 },
 "description":"Dépôts holocènes en Pays de la Loire et en Poitou-Charentes",

"image":"http://external.opengis.org/twiki_public/pub/HydrologyDWG/05-08May2014--GW2IEMeeting18/GW2IE_Vienna2014_Photo2.JPG",
 "label":{
   "@language":"fr",
   "@value":"Dépôts holocènes en Pays de la Loire et en Poitou-Charentes"
 },
 "isPartOf":[
   {
     "@id":"SciObj.7289",
     "@type":"gw:GW_HydrogeoUnit",
     "name":"Grand domaine hydrogéologique des formations sableuses du littor ..."
   }
 ],
 "subjectOf":[
   {
     "@type":"foaf:Document",
     "primaryTopic":"gw:GW_Aquifer",
     "conformsTo":"https://www.opengis.net/def/gwml2",
     "provider":"https://data.geoscience.fr",
     "format":[
       "https://w3id.org/mediatype/application/ld+json",
       "https://w3id.org/mediatype/application/gml+xml"
     ]
   },
   {
     "@type":"foaf:Document",
     "primaryTopic":"sa_saq:EntiteHydroGeol",
     "conformsTo":"http://xml.sandre.eaufrance.fr/saq/2.1/sandre_fmt_xml_saq.xsd",
     "provider":"https://data.geoscience.fr",
     "format":[
       "https://w3id.org/mediatype/application/pdf",
       "https://w3id.org/mediatype/text/html"
     ]
   },
```

```
    {
      "@type":"foaf:Document",
      "primaryTopic":"sa_saq:EntiteHydroGeol",
      "conformsTo":"http://xml.sandre.eaufrance.fr/saq/2.1/sandre_fmt_xml_saq.xsd",
      "provider":"http://www.sandre.eaufrance.fr/",
      "format":[
        "https://w3id.org/mediatype/text/html",
        "https://w3id.org/mediatype/application/gml+xml"
      ]
    }
  ],
  "spatialCoverage":{
    "@type":"Place",
    "geo":{
      "@type":"GeoShape",
      "polygon":"POLYGON ((-1.0905525921863841 46.06034602009783,
-1.534362396360801 46.20266963112101, -1.534871357834444 46.20299206294537,
-2.1409967857081305 46.817742615004946, -2.3018854628323537 46.9897619913443,
-2.5275606084778284 47.29796513991588, -2.5581048194823914 47.373511151470254,
-2.4572579037060227 47.44951499829384, -2.030042855895151 47.643345550015795,
-2.0290888594076546 47.64344529625211, -1.998133682393931 47.63768186458442,
-1.9968870186684125 47.63722568787524, -1.5065592794658917 47.41425424995508,
-1.5029187405526274 47.41250188966548, -1.4599824362934657 47.38723482985455,
-0.3635629979426332 46.272519618976645, -0.3635625589329876 46.271316402862126,
-0.3635619725268311 46.260383702833316, -1.0591610961519313 46.06437816285927,
-1.0797290188160373 46.06114722677742, -1.0820222819905323 46.06085408712582,
-1.08588382415525 46.0603798228799, -1.0905525921863841 46.06034602009783))"
    }
  }
}
```


Listing 1.0 - Aquifers JSON-LD output example.

In order to achieve this goal it has been necessary to extend the module functionalities to include a filtering capabilities support and to intervene in the backwards mapping filtering to provide consistent results when querying the data.

Using the filtering capabilities

Filtering capabilities support allows to define CQL filters inside the template to define whether an array, an object or a value should be present in the final output.

The rules to follow to use the functionality are the following:

- In the array and object cases the filter syntax expects a *$filter* key followed by an attribute with the filter to evaluate, for example, we want the *isPartOf* to only contain the relations matching the relation type *dct:isPartOf*:

```
(...)
"isPartOf":[
  {
    "$source":"ldg:child/ldg:SciObjRelation",
    "$filter":"xpath('ldg:relationType/@xlink:title') = 'dct:isPartOf'"
  },
(...)
```

- In the attribute case, instead, the filter is being specified inside the value as `$filter{...}`, followed by the CQL expression, or by the static content, with a comma separating the two eg.

```
(...)
"subjectOf":[
   {
     "$source":"ldg:subjectOf/ldg:DataResource"
   },
   {
     "@id":"${ldg:URI}",
     "@type":"foaf:Document",
     "primaryTopic":"$filter{xpath(
                    'ldg:child/ldg:SciObj/ldg:primaryTopic')
                    ='gw:GW_Aquifer'},
                    ${ldg:child/ldg:SciObj/ldg:primaryTopic},
(...)
```

The filter evaluation is handled in the following way:

- if a `$filter:cql` attribute is present after the `$source` attribute in an array or an object:
  - in the array case, each array element will be included in the output only if the condition in the filter is matched, otherwise it will be skipped;
  - in the object case, the entire object will be included in the output only if the condition in the filter is matched, otherwise the object will be skipped;

- if a `$filter{cql}` is present inside an attribute value before the expression or the static content, separated by it from a comma

    - in case of an expression attribute, the result of the expression will be included in the output if the filter condition is true;
    - in case of a static content attribute, the static content will be included in the output if the filter condition is true.
    - in case the expression is not matched, the content, static or dynamic, will not be set, resulting in the attribute being skipped.

For the purpose of the aquifer use case this allowed to define filters for the **isPartOf**, including elements in the proper list according to the relation type value (see the JSON-LD template) .

As mentioned above the filtering support is available not only for arrays but also for objects and single attributes. Below some examples.
Since the aquifer use case doesn't present any complex JSON object in the template, for this case the Borehole template will be used as an example:

- JSON Object: the eposb:referenceElevation object will be included in the final output only if the filter condition is met (elevation > 0).

```
...
 "eposb:referenceElevation": {
   "$source": "eposb:referenceElevation",
   "$filter":"xpath('eposb:Elevation/eposb:elevation') > 0 ",
   "@type": "eposb:Elevation",
   "eposb:elevation": "${eposb:Elevation/eposb:elevation}",
   "mdex:verticalCRS": "http://www.opengis.net/def/crs/EPSG/0/5720",
   "om:uom": "http://qudt.org/vocab/unit/M",
   "eposb:elevationMeasurementMethod": {
     "@id":
"${eposb:Elevation/eposb:elevationMeasurementMethod/@xlink:href}",
     "name": "DGPS as Dummy"
   },
   "eposb:elevationType": {
     "@id": "https://data.geoscience.earth/ncl/ElevationType/groundSurface",
     "name": "Ground surface"
   }
 }
...
```

- JSON Attribute (dynamic): the dynamic attribute primaryTopic will be encoded only if it equals “gw:GW_Aquifer”. The same condition has been applied to the attribute staticValue

```
...
"subjectOf":[
       {
         "$source":"ldg:relation/ldg:SciObjRelation",
         "$filter":"xpath('ldg:relationType') = 'dct:hasPart'"
       },
       {
         "@id":"${ldg:child/ldg:SciObj/ldg:uri}",
         "@type":"foaf:Document",
         "primaryTopic":"
          $filter{xpath('ldg:child/ldg:SciObj/ldg:primaryTopic')
          ='gw:GW_Aquifer'},
          ${ldg:child/ldg:SciObj/ldg:primaryTopic}",
         "staticValue":
           "$filter{xpath('ldg:child/ldg:SciObj/ldg:primaryTopic')
           ='gw:GW_Aquifer'}, some optional static value",
         "format":[
           "application/ld+json"
         ],
         "conformsTo":
          "
          ${ldg:child/ldg:SciObj/ldg:conformsTo}",
         "provider":"${ldg:child/ldg:SciObj/ldg:provider}"
       }
     ]
...
```

Requesting using the backwards filtering support

Backwards mapping filtering support has required some modifications in order to maintain consistency when filters are involved in a template: if the template path specified in the filter points to an attribute with a filter or to an attribute inside an array or an object with a filter, the filter will be concatenated with an AND condition to the filter specified through the backward mapping filtering support.

Consider the following CQL filter: `features.isPartOf.name='Sables du Quaternaire en Artois-Picardie'`. Behind the scene GeoServer will concatenate to this filter the template filter, sending to the store the following condition:

```
ldg:SciObj/ldg:relation/ldg:relationType = 'dct:isPartOf'
AND
ldg:SciObj/ldg:relation/ldg:SciObjRelation/ldg:child/ldg:SciObj/gml:name
```

This behaviour is transparent for the user which still needs to simply produce a valid template path in order to take advantage of the backwards mapping support.

Using env parameterization

Env parametrization has been added in order to manipulate the template configuration on the fly via the **env** query parameter.
Env parametrization can be used to manipulate attributes' names, static value, dynamic value as well as portion of expression i.e. xpaths.

Consider the following aquifer template modified to have some *env* functions:

```
{
 "@context":[
   "https://opengeospatial.github.io/ELFIE/contexts/elfie-2/elf-index.jsonld",
   "https://opengeospatial.github.io/ELFIE/contexts/elfie-2/gwml2.jsonld",
   {
     "gsp":"http://www.opengis.net/ont/geosparql#",
     "sf":"http://www.opengis.net/ont/sf#",
     "schema":"https://schema.org/",
     "dc":"http://purl.org/dc/terms/",
     "wkt":"gsp:asWKT",
     "Feature":"gsp:Feature",
     "geometry":"gsp:hasGeometry",
     "polygon":"sf:polygon",
     "primaryTopic":"schema:primaryTopic",
     "features":{
       "@container":"@set",
       "@id":"schema:hasPart"
     }
   }
 ],
 "type":"FeatureCollection",
 "features":[
   {
     "$source":"$${strConcat('ldg:', env('rootSource','SciObj'))}"
   },
   {
     "@id":"${@id}",
     "@type":[
```

```
        "Feature",
        "gw:GW_Aquifer"
      ],
      "name":{
        "@language":"$${env('language','fr')}",
        "@value":"${gml:name}"
      },
      "description":"${gml:description}",

"$${env('image','image')}":"http://external.opengis.org/twiki_public/pub/HydrologyD
WG/05-08May2014--GW2IEMeeting18/GW2IE_Vienna2014_Photo2.JPG",
      "label":{
        "@language":"fr",
        "@value":"${gml:name}"
      },
      "isPartOf":[
        {
          "$source":"ldg:child/ldg:SciObjRelation",
          "$filter":"xpath('ldg:relationType/@xlink:title') = 'dct:isPartOf' AND
                     env('isPartOf','true') = 'true'"
        },
        {
          "@id":"${ldg:child/ldg:SciObj/@id}",
          "@type":"gw:GW_HydrogeoUnit",
          "name":"${ldg:child/ldg:SciObj/gml:name}"
        }
      ],
      "subjectOf":[
        {
          "$source":"ldg:subjectOf/ldg:DataResource"
        },
        {
          "@id":"${ldg:URI}",
          "@type":"foaf:Document",
          "primaryTopic":"${ldg:primaryTopic}",
          "conformsTo":"${ldg:conformsTo}",
          "provider":"${ldg:provider}",
          "format":[
            {
              "$source":"ldg:format/ldg:Format"
            },
            "${ldg:uri}"
          ]
        }
      ],
      "spatialCoverage":{
        "@type":"Place",
        "geo":{
          "@type":"GeoShape",
          "polygon":"$${toWKT(xpath('ldg:spatialCoverage'))}"
```

```
        }
      }
    }
  ]
}
```

Listing 2.0 JSON-LD template with env parametrization.

Each env function has a name and a default value that can be modified using the env queryParameter in a getFeature request. See the below list:

- The *@language* attribute value is defined through an env function having as default value *fr* (`"@language":"$${env('language','fr')}"`). To change the value an example *env* parameter could be **`env=language:en.`**

- The image attribute name is defined through the expression `"$${env('image','image')}",` to modify it to another key, an example *env* parameter could be **`env=image:picture`**.

- The *isPartOf* filter has an *AND* condition with an *env* parameter `"$filter":"xpath('ldg:relationType/@xlink:title') = 'dct:isPartOf' AND env('isPartOf','true') = 'true'"`. To avoid encoding the *isPartOf* array an example *env* parameter would be **`env=isPartOf:false.`**

- It is possible to manipulate all the above values at once in a single request by specifying an *env* request parameter where each *name:value* is separeted from the next by "**;**" eg. **env=languge:en;image:picture;isPartOf:false**.

Validating the JSON-LD response document

The template module provides a validation on the template, trying to resolve all the dynamic attributes (xpath,CQL,sources) against the featureType.
In the context of the JSON-LD output format has also been provided a semantic validation over the features output.
The semantic validation uses the @context object as a kind of schema for the features, and takes advantage of the java implementation for JSON-LD-API 1.0 specifications.

Validation process

The validation process is carried out in three steps using two of the algorithms provided by the JSON-LD api, namely the expansion and the compaction algorithms.

1. The expansion algorithm is executed against the JSON-LD output, expanding each features' attribute name to IRIs, removing those with no reference in the **@context** and the **@context** itself.
2. The compaction algorithm is then executed over the expanded output, putting back the **@context** and shortens to the terms the expanded attribute names, as before the expansion algorithm was executed, thereby producing a JSON-LD document suitable to be compared with the original one.
3. A comparison between the original and the compacted outputs is then executed: if there are differences due missing attributes in the worked document, it means that the expansion algorithm wasn't able to expand all the attributes name to IRIs due missing definition in the **@context**, thus the validation fails.
4. A RuntimeException is thrown with a failure message, listing the attributes that are missing a reference in the **@context**.

In order to trigger the validation a query parameter needs to be provided in the getFeatures request namely `validation=true:`

```
http://brgm-dev.geo-solutions.it/geoserver/ogc/features/collections/ldg:SciObj/items?f=application%2Fld%2Bjson&limit=50&validation=true
```

In case of a failure an error messsage is produced with a message that looks like the followings:

- OGCAPI

```
{
 "code": "NoApplicableCode",
 "description": "org.geoserver.platform.ServiceException:
                 java.lang.RuntimeException: Validation failed.
                 Unable to resolve the following fields
                 against the @context: polygon,isPartOf"
}
```

- WFS

```
<ServiceExceptionReport
       version="1.2.0"
       xmlns="http://www.opengis.net/ogc"
       xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
                          xsi:schemaLocation="http://www.opengis.net/ogc
http://schemas.opengis.net/wfs/1.0.0/OGC-exception.xsd">
   <ServiceException>
     org.geoserver.platform.ServiceException: Validation failed.
     Unable to resolve the following
     fields against the @context:
     polygon,isPartOf
   </ServiceException>
</ServiceExceptionReport>
```

Conclusion

We were able to implement the aquifers use case; a few example requests targeting our DEV environment are available in *Annex A* (see Example requests). We went beyond the original scope by also implementing a JSON-LD validation process that can be requested on demand for each request.

The implemented solution is compatible with both WFS 2.0 and OGC Features - API, we also strive to implement a generic templating approach that will allow us to seamlessly add support for other encoding formats like GeoJSON and GM

Foreseen enhancement

UI and Rest API

Currently all the templates are configured manually by directly editing the data directory. This error prone and go against the best development practices of GeoServer.

The UI could be similar to the one used for sld, providing:

- A text box where it is possible to edit the template.
- Template upload.
- Copy of the template from an existing one.

OGCAPI Complex Features HTML representation

At the moment the html representation for complex features in the ogcapi module doesn't work, but an html representation of complex features is crucial in order to include the features' JSON-LD representation in their html encoding, to facilitate search engines that scrape the `ogcapi`.

GML templating

The templating mechanism is currently supported for JSON-LD and geoJSON output formats.
A GML support could also be provided including all the templating aspects currently implemented: referencing simple or complex features attributes, filtering and env parametrizion support, support for backwards mapping.
GML templating should be based on an XML templating allowing thus to support also HTML templating

Multiple FeatureType request for JSON-LD

At the moment the JSON-LD output format supports single featureType request only. This is due the fact that there is no strategy to handle multiple different JSON-LD contexts in the template. A possible solution for this would be to merge the context of each featureTypes defined in their respective template before starting the features encoding.

Annexes

Annex A

JSON-LD template.

```
{
 "@context":[
   "https://opengeospatial.github.io/ELFIE/contexts/elfie-2/elf-index.jsonld",
   "https://opengeospatial.github.io/ELFIE/contexts/elfie-2/gwml2.jsonld",
   {
     "gsp":"http://www.opengis.net/ont/geosparql#",
     "sf":"http://www.opengis.net/ont/sf#",
     "schema":"https://schema.org/",
     "dc":"http://purl.org/dc/terms/",
     "wkt":"gsp:asWKT",
     "Feature":"gsp:Feature",
     "geometry":"gsp:hasGeometry",
     "polygon":"sf:polygon",
     "primaryTopic":"schema:primaryTopic",
     "features":{
       "@container":"@set",
       "@id":"schema:hasPart"
     }
   }
 ],
 "type":"FeatureCollection",
 "features":[
   {
     "$source":"$${strConcat('ldg:', env('rootSource','SciObj'))}"
   },
   {
     "@id":"${@id}",
     "@type":[
       "Feature",
       "gw:GW_Aquifer"
     ],
     "name":{
       "@language":"$${env('language','fr')}",
       "@value":"${gml:name}"
     },
     "description":"${gml:description}",

"image":"http://external.opengis.org/twiki_public/pub/HydrologyDWG/05-08May2014--GW
2IEMeeting18/GW2IE_Vienna2014_Photo2.JPG",
     "label":{
       "@language":"fr",
```

```
        "@value":"${gml:name}"
      },
      "isPartOf":[
        {
          "$source":"ldg:child/ldg:SciObjRelation",
          "$filter":"xpath('ldg:relationType/@xlink:title') = 'dct:isPartOf'"
        },
        {
          "@id":"${ldg:child/ldg:SciObj/@id}",
          "@type":"gw:GW_HydrogeoUnit",
          "name":"${ldg:child/ldg:SciObj/gml:name}"
        }
      ],
      "subjectOf":[
        {
          "$source":"ldg:subjectOf/ldg:DataResource"
        },
        {
          "@id":"${ldg:URI}",
          "@type":"foaf:Document",
          "primaryTopic":"${ldg:primaryTopic}",
          "conformsTo":"${ldg:conformsTo}",
          "provider":"${ldg:provider}",
          "format":[
            {
              "$source":"ldg:format/ldg:Format"
            },
            "${ldg:uri}"
          ]
        }
      ],
      "spatialCoverage":{
        "@type":"Place",
        "geo":{
          "@type":"GeoShape",
          "polygon":"$${toWKT(xpath('ldg:spatialCoverage'))}"
        }
      }
    }
  ]
}
```

Example requests

The following are example requests aimed at showing the newly implemented capabilities.

1. Requests targeting the SciObj type as JSON-LD.

    a. WFS 2.0:
    http://brgm-dev.geo-solutions.it/geoserver/ldg/ows?service=WFS&version2.0.0&request=GetFeature&typeNames=ldg:SciObj&count=50&outputFormat=application%2Fld%2Bjson
    b. OGCAPI:
    http://brgm-dev.geo-solutions.it/geoserver/ogc/features/collections/ldg:SciObj/items?f=application%2Fld%2Bjson&limit=50

2. Requests targeting the SciObj type using a CQL filter pointing to a template attribute expression inside a filtered array (`features.isPartOf.name='Sables du Quaternaire en Artois-Picardie'`):

    a. WFS 2.0:
    http://brgm-dev.geo-solutions.it/geoserver/ldg/ows?service=WFS&version=2.0.0&request=GetFeature&typeNames=ldg:SciObj&count=50&outputFormat=application%2Fld%2Bjson&cql_filter=features.isPartOf.name='Sables du Quaternaire en Artois-Picardie'

    b. OGCAPI**:**
    http://brgm-dev.geo-solutions.it/geoserver/ogc/features/collections/ldg:SciObj/items?f=application%2Fld%2Bjson&limit=50&filter-lang=cql-text&filter=features.isPartOf.name='Sables du Quaternaire en Artois-Picardie'

3. Request targeting the SciObj type using an env parameter to block the display of isPartOf array (`env=isPartOf:false`).

    a. WFS 2.0:
    http://brgm-dev.geo-solutions.it/geoserver/ldg/ows?service=WFS&version2.0.0&request=GetFeature&typeNames=ldg%3ASciObj&count=50&outputFormat=application%2Fld%2Bjson&env=isPartOf:false

    b. OGCAPI:

http://brgm-dev.geo-solutions.it/geoserver/ogc/features/collections/ldg:SciObj/items?f=application%2Fld%2Bjson&limit=50&env=isPartOf:false

Annex B

XSD

```
<?xml version="1.0" encoding="UTF-8"?>
<schema xmlns="http://www.w3.org/2001/XMLSchema"
xmlns:gmd="http://www.isotc211.org/2005/gmd"
xmlns:gml="http://www.opengis.net/gml/3.2"
xmlns:ldg="https://data.geoscience.fr/def/LinkedDataGazetteer"
targetNamespace="https://data.geoscience.fr/def/LinkedDataGazetteer"
elementFormDefault="qualified" version="0.2">
   <import namespace="http://www.isotc211.org/2005/gmd"
schemaLocation="http://schemas.opengis.net/iso/19139/20070417/gmd/gmd.xsd" />
   <import namespace="http://www.opengis.net/gml/3.2"
schemaLocation="http://schemas.opengis.net/gml/3.2.1/gml.xsd" />
   <!-- XML Schema document created by ShapeChange - http://shapechange.net/ -->
   <element name="Format" type="ldg:FormatType" substitutionGroup="gml:AbstractGML"
/>
   <complexType name="FormatType">
      <complexContent>
         <extension base="AbstractGMLType">
            <sequence>
               <element name="uri" type="anyURI" />
            </sequence>
         </extension>
      </complexContent>
   </complexType>
   <complexType name="FormatPropertyType">
      <sequence minOccurs="0">
         <element ref="ldg:Format" />
         <attributeGroup ref="gml:AssociationAttributeGroup" />
      </sequence>
   </complexType>
   <element name="DataResource" type="ldg:DataResourceType"
substitutionGroup="gml:AbstractGML">
      <annotation>
         <documentation>DIR in SELFIE lingua franca. DataResource providing one
or more representations of the Scientific Object.</documentation>
         <documentation>20200827-S.Grellet : /ldg:DataResource/ldg:URI
typed</documentation>
      </annotation>
   </element>
```

```
   <complexType name="DataResourceType">
       <complexContent>
           <extension base="gml:AbstractGMLType">
               <sequence>
                   <element name="URI" type="anyURI">
                       <annotation>
                           <documentation>accessURL of the data
                                          Resource.
                           </documentation>
                       </annotation>
                   </element>
                   <element name="primaryTopic" type="string">
                       <annotation>
                           <documentation>primaryTopic of the
                                          DataResource.
                           </documentation>
                       </annotation>
                   </element>
                   <element name="conformsTo" type="string">
                       <annotation>
                           <documentation>Reference to (URI) the specification the
DataResource is conformTo.</documentation>
                       </annotation>
                   </element>
                   <element name="provider" type="string">
                       <annotation>
                           <documentation>URI (ideally) identifying the provider of
the DataResource.</documentation>
                       </annotation>
                   </element>
                   <element name="format" type="ldg:FormatPropertyType"
minOccurs="0" maxOccurs="unbounded" />
               </sequence>
           </extension>
       </complexContent>
   </complexType>
   <complexType name="DataResourcePropertyType">
       <sequence minOccurs="0">
           <element ref="ldg:DataResource" />
       </sequence>
       <attributeGroup ref="gml:AssociationAttributeGroup" />
       <attributeGroup ref="gml:OwnershipAttributeGroup" />
   </complexType>
   <element name="SciObj" type="ldg:SciObjType"
substitutionGroup="gml:AbstractFeature">
       <annotation>
           <documentation>NIR in SELFIE lingua franca (/id). The Scientific Object
we are describing.</documentation>
       </annotation>
   </element>
```

```
   <complexType name="SciObjType">
       <complexContent>
           <extension base="gml:AbstractFeatureType">
               <sequence>
                   <element name="spatialCoverage" type="gml:GeometryPropertyType">
                       <annotation>
                           <documentation>Convex hull of the Scientif Object
corresponding to its spatial coverage.</documentation>
                       </annotation>
                   </element>
                   <element name="type" type="gml:ReferenceType"
maxOccurs="unbounded" />
                   <element name="parent" type="ldg:SciObjRelationPropertyType">
                       <annotation>
                           <appinfo>
                               <reversePropertyName
xmlns="http://www.opengis.net/gml/3.2">ldg:child</reversePropertyName>
                           </appinfo>
                       </annotation>
                   </element>
                   <element name="child" type="ldg:SciObjRelationPropertyType"
minOccurs="0" maxOccurs="unbounded">
                       <annotation>
                           <appinfo>
                               <reversePropertyName
xmlns="http://www.opengis.net/gml/3.2">ldg:parent</reversePropertyName>
                           </appinfo>
                       </annotation>
                   </element>
                   <element name="subjectOf" type="ldg:DataResourcePropertyType"
maxOccurs="unbounded" />
               </sequence>
           </extension>
       </complexContent>
   </complexType>
   <complexType name="SciObjPropertyType">
       <sequence minOccurs="0">
           <element ref="ldg:SciObj" />
       </sequence>
       <attributeGroup ref="gml:AssociationAttributeGroup" />
       <attributeGroup ref="gml:OwnershipAttributeGroup" />
   </complexType>
   <element name="SciObjRelation" type="ldg:SciObjRelationType"
substitutionGroup="gml:AbstractGML" />
   <complexType name="SciObjRelationType">
       <complexContent>
           <extension base="gml:AbstractGMLType">
               <sequence>
                   <element name="relationType" type="gml:ReferenceType">
                       <annotation>
```

```
                        <documentation>isPartOf, relatedTo</documentation>
                    </annotation>
                </element>
                <element name="parent" type="ldg:SciObjPropertyType">
                    <annotation>
                        <appinfo>
                            <reversePropertyName
xmlns="http://www.opengis.net/gml/3.2">ldg:child</reversePropertyName>
                        </appinfo>
                    </annotation>
                </element>
                <element name="child" type="ldg:SciObjPropertyType"
minOccurs="0" maxOccurs="unbounded">
                    <annotation>
                        <appinfo>
                            <reversePropertyName
xmlns="http://www.opengis.net/gml/3.2">ldg:parent</reversePropertyName>
                        </appinfo>
                    </annotation>
                </element>
            </sequence>
        </extension>
    </complexContent>
  </complexType>
  <complexType name="SciObjRelationPropertyType">
    <sequence minOccurs="0">
        <element ref="ldg:SciObjRelation" />
    </sequence>
    <attributeGroup ref="gml:AssociationAttributeGroup" />
    <attributeGroup ref="gml:OwnershipAttributeGroup" />
  </complexType>
</schema>
```

MAPPINGS

```
<as:AppSchemaDataAccess xmlns:as="http://www.geotools.org/app-schema"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.geotools.org/app-schema AppSchemaDataAccess.xsd"
xmlns:ldg="https://data.geoscience.fr/def/LinkedDataGazetteer">
  <namespaces>
    <Namespace>
      <prefix>gml</prefix>
      <uri>http://www.opengis.net/gml/3.2</uri>
    </Namespace>
    <Namespace>
```

```
            <prefix>ldg</prefix>
            <uri>https://data.geoscience.fr/def/LinkedDataGazetteer</uri>
        </Namespace>
        <Namespace>
            <prefix>xlink</prefix>
            <uri>http://www.w3.org/1999/xlink</uri>
        </Namespace>
        <Namespace>
            <prefix>xsi</prefix>
            <uri>http://www.w3.org/2001/XMLSchema-instance</uri>
        </Namespace>
    </namespaces>
    <sourceDataStores>
        <DataStore>
            <id>datastore</id>
            <parameters>
                <Parameter>
                    <name>dbtype</name>
                    <value>postgis</value>
                </Parameter>
                <Parameter>
                    <name>host</name>
                    <value>localhost</value>
                </Parameter>
                <Parameter>
                    <name>port</name>
                    <value>5432</value>
                </Parameter>
                <Parameter>
                    <name>database</name>
                    <value>gazetturi</value>
                </Parameter>
                <Parameter>
                    <name>user</name>
                    <value>postgres</value>
                </Parameter>
                <Parameter>
                    <name>passwd</name>
                    <value>postgres</value>
                </Parameter>
                <Parameter>
                    <name>schema</name>
                    <value>linkeddatagazetteer</value>
                </Parameter>
                <Parameter>
                    <name>Expose primary keys</name>
                    <value>true</value>
                </Parameter>
            </parameters>
        </DataStore>
```

```
</sourceDataStores>
<targetTypes>
    <FeatureType>
        <schemaUri>LinkedDataGazetter.xsd</schemaUri>
    </FeatureType>
</targetTypes>
<typeMappings>
    <FeatureTypeMapping>
        <mappingName>ldg:SciObj</mappingName>
        <sourceDataStore>datastore</sourceDataStore>
        <sourceType>sciobj</sourceType>
        <targetElement>ldg:SciObj</targetElement>
        <attributeMappings>
            <AttributeMapping>
                <targetAttribute>ldg:SciObj</targetAttribute>
                <idExpression>
                    <OCQL>Concatenate('SciObj.', sciobj_id)</OCQL>
                </idExpression>
            </AttributeMapping>
            <AttributeMapping>
                <targetAttribute>gml:name</targetAttribute>
                <sourceExpression>
                    <OCQL>name</OCQL>
                </sourceExpression>
            </AttributeMapping>
            <AttributeMapping>
                <targetAttribute>gml:description</targetAttribute>
                <sourceExpression>
                    <OCQL>description</OCQL>
                </sourceExpression>
            </AttributeMapping>
            <AttributeMapping>
                <targetAttribute>ldg:spatialCoverage</targetAttribute>
                <ClientProperty>
                    <name>codeSpace</name>
                    <value>'http://www.ietf.org/rfc/rfc2616'</value>
                </ClientProperty>
                <sourceExpression>
                    <OCQL>spatialcoverage</OCQL>
                </sourceExpression>
            </AttributeMapping>
            <AttributeMapping>
                <targetAttribute>ldg:child</targetAttribute>
                <sourceExpression>
                    <OCQL>sciobj_id</OCQL>
                    <linkElement>'SciObjRelationType'</linkElement>
                    <linkField>FEATURE_LINK[1]</linkField>
                </sourceExpression>
                <isMultiple>true</isMultiple>
            </AttributeMapping>
```

```
            <AttributeMapping>
                <targetAttribute>ldg:subjectOf</targetAttribute>
                <sourceExpression>
                    <OCQL>sciobj_id</OCQL>
                    <linkElement>'ldg:DataResource'</linkElement>
                    <linkField>FEATURE_LINK[3]</linkField>
                </sourceExpression>
                <isMultiple>true</isMultiple>
            </AttributeMapping>
        </attributeMappings>
    </FeatureTypeMapping>
    <FeatureTypeMapping>
        <mappingName>ldg:DataResource</mappingName>
        <sourceDataStore>datastore</sourceDataStore>
        <sourceType>dataresource</sourceType>
        <targetElement>ldg:DataResource</targetElement>
        <attributeMappings>
            <AttributeMapping>
                <targetAttribute>ldg:primaryTopic</targetAttribute>
                <sourceExpression>
                    <OCQL>primarytopic</OCQL>
                </sourceExpression>
            </AttributeMapping>
            <AttributeMapping>
                <targetAttribute>ldg:provider</targetAttribute>
                <sourceExpression>
                    <OCQL>provider</OCQL>
                </sourceExpression>
            </AttributeMapping>
            <AttributeMapping>
                <targetAttribute>ldg:conformsTo</targetAttribute>
                <sourceExpression>
                    <OCQL>conformsto</OCQL>
                </sourceExpression>
            </AttributeMapping>
            <AttributeMapping>
                <targetAttribute>FEATURE_LINK[3]</targetAttribute>
                <sourceExpression>
                    <OCQL>sciobj_id</OCQL>
                </sourceExpression>
            </AttributeMapping>
            <AttributeMapping>
                <targetAttribute>ldg:format</targetAttribute>
                <sourceExpression>
                    <OCQL>dataresource_id</OCQL>
                    <linkElement>'ldg:Format'</linkElement>
                    <linkField>FEATURE_LINK[4]</linkField>
                </sourceExpression>
                <isMultiple>true</isMultiple>
            </AttributeMapping>
```

```
            </attributeMappings>
        </FeatureTypeMapping>
        <FeatureTypeMapping>
            <mappingName>ldg:Format</mappingName>
            <sourceDataStore>datastore</sourceDataStore>
            <sourceType>dataresourceformaturi</sourceType>
            <targetElement>ldg:Format</targetElement>
            <attributeMappings>
                <AttributeMapping>
                    <targetAttribute>ldg:uri</targetAttribute>
                    <sourceExpression>
                        <OCQL>formaturi</OCQL>
                    </sourceExpression>
                </AttributeMapping>
                <AttributeMapping>
                    <targetAttribute>FEATURE_LINK[4]</targetAttribute>
                    <sourceExpression>
                        <OCQL>dataresource_id</OCQL>
                    </sourceExpression>
                    <isMultiple>true</isMultiple>
                </AttributeMapping>
            </attributeMappings>
        </FeatureTypeMapping>
        <FeatureTypeMapping>
            <mappingName>SciObjRelationType</mappingName>
            <sourceDataStore>datastore</sourceDataStore>
            <sourceType>sciobjrelation</sourceType>
            <targetElement>ldg:SciObjRelation</targetElement>
            <attributeMappings>
                <AttributeMapping>
                    <targetAttribute>ldg:SciObjRelation</targetAttribute>
                    <idExpression>
                        <OCQL>Concatenate('SciObjRelation.',
sciobjrelation_id)</OCQL>
                    </idExpression>
                </AttributeMapping>
                <AttributeMapping>
                    <targetAttribute>ldg:relationType</targetAttribute>
                    <encodeIfEmpty>true</encodeIfEmpty>
                    <ClientProperty>
                        <name>xlink:title</name>
                        <value>relationtype</value>
                    </ClientProperty>
                </AttributeMapping>
                <AttributeMapping>
                    <targetAttribute>FEATURE_LINK[1]</targetAttribute>
                    <sourceExpression>
                        <OCQL>parent_id</OCQL>
                    </sourceExpression>
                </AttributeMapping>
```

```
            <AttributeMapping>
                <targetAttribute>ldg:child</targetAttribute>
                <sourceExpression>
                    <OCQL>child_id</OCQL>
                    <linkElement>'ldg:SciObjChild'</linkElement>
                    <linkField>FEATURE_LINK[2]</linkField>
                </sourceExpression>
                <isMultiple>true</isMultiple>
            </AttributeMapping>
        </attributeMappings>
    </FeatureTypeMapping>
    <FeatureTypeMapping>
        <mappingName>ldg:SciObjChild</mappingName>
        <sourceDataStore>datastore</sourceDataStore>
        <sourceType>sciobj</sourceType>
        <targetElement>ldg:SciObj</targetElement>
        <attributeMappings>
            <AttributeMapping>
                <targetAttribute>ldg:SciObj</targetAttribute>
                <idExpression>
                    <OCQL>Concatenate('SciObj.', sciobj_id)</OCQL>
                </idExpression>
            </AttributeMapping>
            <AttributeMapping>
                <targetAttribute>gml:name</targetAttribute>
                <sourceExpression>
                    <OCQL>name</OCQL>
                </sourceExpression>
            </AttributeMapping>
            <AttributeMapping>
                <targetAttribute>gml:description</targetAttribute>
                <sourceExpression>
                    <OCQL>description</OCQL>
                </sourceExpression>
            </AttributeMapping>
            <AttributeMapping>
                <targetAttribute>FEATURE_LINK[2]</targetAttribute>
                <sourceExpression>
                    <OCQL>sciobj_id</OCQL>
                </sourceExpression>
            </AttributeMapping>
        </attributeMappings>
    </FeatureTypeMapping>
  </typeMappings>
</as:AppSchemaDataAccess>
```

In order the map the many to many relationship between the DataResourceType and the FormatType, a denormalized view has been created using the following query:

```
create view linkeddatagazetteer.dataresourceformatURI as select dataresource_id,
uri as formaturi from  linkeddatagazetteer.dataresourceformat drf join
linkeddatagazetteer.reg_format
on drf.format_id=reg_format.reg_format_id order by dataresource_id
```

SQL query 1.0 - Denormalized view