Publishing Reference Geodata on the Web: Opportunities and Challenges for IGN France

Ghislain A. Atemezing[1], Nathalie Abadie[2],
Raphaël Troncy[1], Bénédicte Bucher[2]

[1] EURECOM, SophiaTech Campus, France,
[2] Université Paris-Est, IGN /SRIG, COGIT, Saint-Mandé

Abstract. The French national mapping agency (IGN) produces several different but complementary geographic vector reference databases delivered in traditional GIS formats. However, linked data users have different expectations and habits, such as the need to browse an entire data catalogue in RDF using the "follow-your-nose" navigation capacity from one graph to another. Besides, traditional GIS data formats are not interoperable with RDF. Yet, all these geographic datasets could be used with benefits on the Web of data, either with direct georeferencing through geographic primitives, or indirect one through postal addresses. In this paper, we aim to contribute to the georeferencing of datasets published on the Web of data by providing such resources for French context. Firstly, we propose two vocabularies designed for representing structured geometries defined with coordinates expressed in any Coordinates Reference System (CRS). Secondly, we reuse these vocabularies and the CRSs' dataset to publish a reference dataset on administrative units that can also be reused for indirect georeferencing purposes. Finally, we also propose two vocabularies for describing geographic feature types. In addition to these resources, we also present a comprehensive workflow for easily publishing geographic data on the Web of data.

Keywords: Ontology Design, Geospatial Data, Linked Data, Georeferencing, Structured Geometry, Coordinate Reference System, data.ign.fr

1 Introduction

The French national mapping agency (IGN) produces several different but complementary geographic vector reference databases (BD TOPO®, BD CARTO®, BD ADRESSE®, etc.). They are structured according to object-oriented application schemas (ISO 19109). As an example, GEOFLA® database contains data on the French administrative units. Their boundaries are described by geometries of type MultiPolygon and their properties such as toponyms, population, legal codes and hierarchical relationships are stored by attributes. All these databases are provided in traditional GIS formats (ESRI shapefiles or GML). As required by the INSPIRE Directive, IGN provides users with a data visualization portal[3].

[3] `http://www.geoportail.gouv.fr/accueil`

However, linked data users have different expectations and habits. They need to browse the entire data catalogue in RDF and wish to have *"follow-your-nose"* navigation possibility from one graph to another. Besides, GIS data formats are not interoperable with RDF. Indeed, many resources published on the Web of data are georeferenced, either directly through geographic coordinates, geometric primitives or indirectly through postal addresses, names of administrative units or points of interest. According to LOD cloud statistics, the properties `geo:long` and `geo:lat` of the W3C vocabulary Geo[4] are respectively used in 530 450 and 530 515 triples within 59 datasets, while 36 datasets reuse classes defined by 6 different vocabularies describing postal addresses[5]. We have also identified more than 80 properties with semantic meaning closed to :locatedIn or :hasLocation.

In this article, we propose to contribute to the georeferencing of datasets published on the Web of data by providing some useful resources. Firstly, we propose two vocabularies designed for representing structured geometries defined with coordinates expressed in any Coordinates Reference System (CRS). A dataset dedicated to the description of CRSs defined and maintained by IGN France is also published and can be reused for direct georeferencing purposes. Secondly, we reuse these vocabularies and the CRSs' dataset to publish a reference dataset on administrative units that can also be reused for indirect georeferencing purposes. Finally, we also propose two vocabularies for describing geographic feature types. In addition to these resources, we also present a comprehensive workflow for easily publishing geographic data on the Web of data.

The remainder of this article is structured as follows. In Section 2, we present some technical considerations on data georeferencing and publishing on the Web of data. The Section 3 describes the vocabularies developed for topographic features and their geometries. We then present the generation and publication of administrative units (Section 4) and the French gazetteer in Section 5. We conclude the paper with some challenges (Section 6). Finally some conclusions are drawn.

2 Georeferencing data and Technical considerations

Georeferencing data either by direct or indirect spatial reference requires some reference datasets that can be used as the spatial frame for anchoring these thematic data. Especially, it requires data on both CRSs and named places, which must be published on the Web of data.

2.1 Direct georeferencing of data on the Web

Modeling direct location information such as coordinates or vector data geometries in RDF still poses some challenges. In [1], we have conducted a survey of the vocabularies used for representing geographical features from vocabularies

[4] `http://www.w3.org/2003/01/geo/wgs84_pos#`
[5] `http://stats.lod2.eu/`

of feature types to vocabularies for geometric primitives which provide ways for representing extents, shapes and boundaries of those features. Most of vocabularies dedicated to geometry representation reuse W3C Geo vocabulary which allows only WGS84 coordinates, such as NeoGeo[6]. With the rise of the Open Data movement, more and more publishers including governments and local authorities are releasing legacy data that are georeferenced using others CRSs. For example, IGN France releases data using different projected CRSs depending on the geographic extent of each dataset. In order to overcome this limitation on CRSs, the vocabulary designed by OGC GeoSPARQL standard does not reuse W3C Geo vocabulary but proposes another class "Point" instead. Geometries of geographical data represented in RDF with the GeoSPARQL vocabulary are represented by literals encoded consistently with other OGC standards. `gsp:wktLiteral` and `gsp:gmlLiteral` are thus respectively derived from Well-Known Text and GML encoding rules. In `wktLiteral` and `gmlLiteral`, the CRS used to define the coordinates of the point is identified by a dereferenceable URI which is explicitly stated at the beginning of the literal. This way of associating coordinate reference systems with geometries has the advantage of being consistent with Linked Data principles: each CRS is identified with a dereferenceable URI. The main drawback is that such literals cannot be easily queried with SPARQL, unless using regular expression-based filters. To overcome this limitation, we propose in the geometry vocabulary presented in Section 3 to associate each geometry to the CRS used by its coordinates with the property `geom:crs`.

2.2 Indirect georeferencing of data on the Web

Modeling indirect location information such as administrative units or named points of interest in RDF is preferably done by identifying such geographic features with URIs and describing them by their properties, so that they can be referenced by other datasets. This is the case in one of the most reused datasets of the Web of data, namely Geonames[7]. However, there are yet very few reference datasets for the French territory on the Web of data. A simple example is the current resource for *Paris* in the French DBpedia[8]. The department's name associated to this resource is a literal named "Paris" and the different arrondissements composing the city are modeled as `skos:Concept` instead of `dbpedia-owl:Place`. Even Geonames data remain very limited, as French administrative units are provided as simple geometries (POINT). The "Official Geographic Code"[9] published by the French Statistical Institute (INSEE) is the most up-to-date and accurate dataset on French administrative units, but unfortunately it contains no geometrical description of their boundaries. This has the consequence of not having a baseline during mapping process for application developers trying to consume specific data coming from France. Datasets

[6] `http://geovocab.org/doc/neogeo/`
[7] `http://sws.geonames.org/`
[8] `http://fr.dbpedia.org/resource/Paris`
[9] `http://rdf.insee.fr/sparql`

describing administrative units, points of interest or postal addresses with their labels and geometries, and identifying these features with URIs could be used with benefits not only for georeferencing other datasets, but also for interlinking datasets georeferenced by direct and indirect location information.

2.3 Publishing French geographic data on the Web

In order to be published on the Web of data, geographic data must be transformed from their traditional GIS formats into RDF. They must be refined using suitable vocabularies which can be either created for that specific purpose or reused thanks to some catalogue such as LOV[10] [9]. Geographic features must be identified by URIs created according to well-defined policies. Licenses must be attached to the datasets. Additionally, data must be interlinked with various datasets already published on the Linked Open Data cloud. All these steps require specific tools and skills, so that only a few geographic datasets have been published yet in RDF by National Mapping Agencies.

The Ordnance Survey Linked Data Platform[11] has published three products as Linked Data : Gazetteer, Code-Point and the administrative geography for Great Britain [3]. They also provide a wide range of APIs for accessing the different datasets. For visualizing, a Linked Data API[12] is used on top of the datasets. Similar initiative was presented in [2] for Spanish geographical datasets. Although the authors use an ontology network for the modeling, it is difficult at the moment to reuse their vocabulary for geometry because it is more specific to their use-case. However, the availability of complex geometry both in OGC standards and in more-structured RDF is interesting and should be adopted for our use case. Regarding tools integrating workflow for dealing with geodata, the GeoKnow stack[13] offers a set of tools to publish and visualize geodata. But GeoKnow stack is more oriented to expert users in Semantic Technologies. That is why we chose the Datalift Platform [8] among other solutions because it includes almost all of the aforementioned functionalities to publish geographic data as Linked Data, and integrates a geographic data converter. Moreover, it can be used with a variety of triple stores, and more important, it is target at lay users.

3 Vocabularies for Geometries and Feature Types

Direct georeferencing of data implies representing coordinates or geometries and associating them to a CRS. This requires vocabularies for geometries and CRSs. Besides, indirect georeferencing of data implies associating them to other data on named places. Preferably, these data on named places should be also georeferenced by coordinates in order to serve as basis for data linking between indirectly

[10] http://lov.okfn.org/dataset/lov/
[11] http://data.ordnancesurvey.co.uk
[12] http://code.google.com/p/elda/
[13] http://stack.linkeddata.org/download/

and directly georeferenced datasets. In this section, we present the vocabularies that we have defined and reused for geographic data publishing.

3.1 A vocabulary for geometries

In [1], we already surveyed numerous vocabularies for representing geographical features and their geometries, either using a literal (e.g. wktLiteral) or a structured representation à la NeoGeo. We concluded the survey with some recommendations for geometry descriptions:

- the distinction of geometry versus feature and a property linking both classes (e.g. for attaching provenance information on how some points of a geometry have been collected),
- the ability to represent structured geometries (e.g. for performing simple spatial queries on the data, even when they are stored in a triple store that do not implement the GeoSPARQL standard),
- the integration of any coordinate reference system (e.g. for allowing projected coordinates for cartographic purposes).

In addition to these recommendations, we also think that the domain of the property used to link a feature to its geometry should be left empty in order to accept links between any type of resource and a geometry. This would be useful for example, to associate a person to the coordinates of their birthplace.

Extending GeoSPARQL vocabulary In order to fulfill these recommendations, we have developed a new vocabulary that re-uses and extends the existing vocabularies for representing geometries, namely:

- `http://www.opengis.net/ont/geosparql#` (prefix `gsp`). This vocabulary provides the basic concepts to represent geographical data such as SpatialObject, Feature or `Geometry`. A Feature is linked to a Geometry via the relation `gsp:hasGeometry`. The geometries are typed strings (`gsp:gmlLiteral` or `gsp:wktLiteral` corresponding respectively to the properties `gsp:asGML` and `gsp:asWKT`). The vocabulary contains also spatial functions.
- `http://www.opengis.net/ont/sf#` (prefix `sf`): This vocabulary is based on the OGC standard Simple Features Access [5]. The class `sf:Geometry` is a subclass of `gsp:Geometry`.

Reusing and extending GeoSPARQL Simple Features vocabulary with structured geometries à la NeoGeo enables us to represent geometries both with GeoSPARQL compliant literals and with structured geometries that can be handled easily with SPARQL. The extension for structured geometries consists in defining a subclass for each class from the `sf` vocabulary, and defining properties to associate its instances with a CRS and coordinates or other suitable geometric primitives. For example, the class `geom:Point` is a subclass of `sf:Point`. An instance of `geom:Point` is associated with exactly one instance of `ignf:CRS` via the property `geom:crs`, and it has exactly one coordinate X and exactly one coordinate Y. It can also have a Z coordinate. The coordinates are `xsd:double` and correspond to the properties `geom:coordX:`, `geom:coordY:` and `geom:coordZ:` respectively. Other complex geometries are also defined, such as Linestrings, LinearRings, Polygons or MultiPolygons. Their definitions are based on the class `geom:Point`. As an example, an instance of `geom:Linestring` is defined as an instance of `geom:PointsList` which is an ordered `rdf:List` of instances of `geom:Point` designated by the property `geom:points`.

We have also defined a property `geom:geometry` with an empty domain. Thus, our proposal defines a more generic class for a POINT with the benefit of choosing the CRS of the underlying data. Figure 1 gives an overview of the relationships between the high level concepts with geometries, CRS and topographic features.

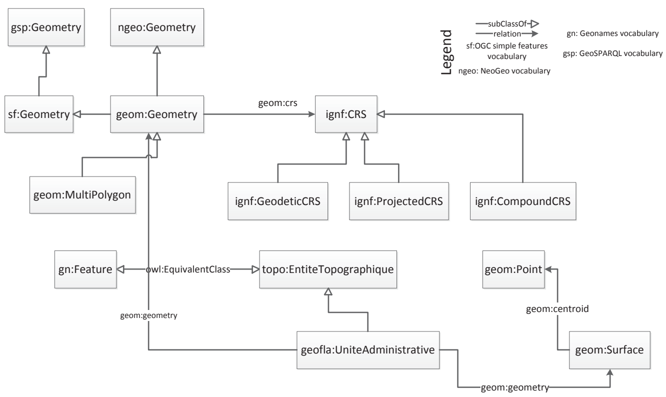

Fig. 1: High level classes of ignf, geom and topo vocabularies; relationships between them and mappings with external vocabularies.

3.2 CRS requirements for the French territory

As explained in Section 2, making explicit the CRS used in a given dataset is a very important issue when dealing with direct location data. This is especially important in the field of geographical information where different CRSs are commonly used due to technical or legal requirements. For INSPIRE Directive, CRS are considered as reference data used for linking thematic data [4], and must be described according to ISO 19111 standard. To be consistent with Linked Data principles, CRS should be identified by URIs, like in OGC proposal. Moreover, as Linked Data users are not always familiar with CRS identifiers commonly used within the geographic information community, URI used to identify CRS should use more intuitive names. Finally, consistently with our goal of contributing to a better georeferencing of data on the French territory, we need an access to the descriptions of all French CRSs, including some deprecated but still used CRSs like “Lambert 1”.

3.3 Identifying and describing CRSs on the Web

In order to fulfill the need for CRS identification and description on the Web, OGC maintains a set of URIs for identifying the most commonly used CRS. While very useful, the main disadvantage of this proposal is that the URIs defined by OGC are not very intuitive for users who are not familiar with Spatial Reference System Identifiers defined by geographic information authorities like OGC or EPSG, such as “4326” (which actually refers to a WGS84 CRS defined by the EPSG). Moreover, many CRS commonly used locally, such as deprecated French projected CRS, are not available in that registry. In addition to OGC proposal, several registries have been proposed by the geographic information community for cataloguing existing CRSs. The EPSG Geodetic Parameter Registry[14] allows querying the Geodetic Parameter Dataset gathered by the EPSG. CRSs can be retrieved by name, by code, by type or by coverage area, and their characteristics are displayed on a HTML form. Unfortunately, there is no direct access to these data through dereferenceable URIs. The Information and Service System for European Coordinate Reference Systems[15] provides an access to ISO 19111 standard-based descriptions of the main European CRSs but has the same limitation as the EPSG registry: access to the descriptions is not allowed by URI, but only through a cartographic interface. `SpatialReference.org` initiative aims at allowing users to use URI-based references to spatial reference systems, including some CRSs defined and maintained by IGN France. Besides, the proposed URL policy is not very intuitive. As an example, this URL identifies the projected system defined by IGN France, Lambert 93: `http://spatialreference.org/ref/sr-org/7527/`. Moreover, the definitions of some deprecated CRSs such as Lambert zone projected CRSs (which are still used in some datasets) seem to be referenced only for the authority EPSG and not for IGNF, which also maintains a registry of CRSs. ISO 19111 standard-based definitions of all CRSs defined and maintained by IGN France are published in an XML file[16]. References to equivalent definitions provided by the EPSG registry are explicitly stated with EPSG SRID. CRSs are identified by URIs using short names instead of numeric codes. For example, `http://registre.ign.fr/ign/IGNF/crs/NTFLAMB2E` is the URI designed for the “Lambert 2 étendu” projected system. Indeed “NTFLAMB2E”

Prefix	URI
geofla	http://data.ign.fr/def/geofla#
geom	http://data.ign.fr/def/geometrie#
ignf	http://data.ign.fr/def/ignf#
rgeofla	http://data.ign.fr/id/geofla/
topo	http://data.ign.fr/def/topo#
rtopo	http://data.ign.fr/id/topo/

Table 1: URI schemes and conventions used for vocabularies and resources .

[14] `http://www.epsg-registry.org/`
[15] `http://www.crs-geo.eu`
[16] `http://librairies.ign.fr/geoportail/resources/IGNF.xml`

is used to identify the projected system "Lambert 2 étendu" which is based on NTF (New French Triangulation) geodetic reference system. Unfortunately, this registry is still in evolution and its URIs are not dereferenceable yet.

As no existing registry fulfilled all our requirements, we have developed a vocabulary[17], inspired from the ISO 19111 schema for CRSs description. Then we have converted IGNF CRSs registry into RDF, and published this dataset on the Web with the Datalift platform[18]. Therefore, the description of the "NTF Lambert 2 étendu" projected CRS can be retrieved at this URL `http://data.ign.fr/id/ignf/crs/NTFLAMB2E`.

3.4 Vocabularies for Geographic Feature Types

Indirect georeferencing of resources on the Web requires reference geographic data on named places and therefore vocabularies for describing feature types and their properties. Therefore, we have chosen to publish a reference dataset on administrative units called GEOFLA®, which is already available in GIS format under an Open Data license. We have also made tests of data conversion and interlinking with another largest dataset on French names places. We have produced and published two vocabularies to describe these datasets, to make sure that all concepts and properties needed would be available. In the GEOFLA® vocabulary[19], 5 classes have been defined: commune, canton, arrondissement, department and region. In the BD TOPO® vocabulary[20] 35 main classes have been defined. They represent the main types of geographic features represented in the BD TOPO® database. In both vocabularies, properties have been defined based on the attributes of their related classes in the databases. The geographic feature types defined as values of attributes "nature" are modeled as instances of `skos:Concept`. SKOS is intensively used to easily group concepts into different schemes (using `skos:hasTopConcept`) and provide semantic relationships (e.g: `skos:broader`, `skos:narrowMatch`) among them. We also provide alignments with Geonames vocabulary, where `topo:Place` is subclass of `gn:S` and `owl:sameAs` linked concepts.[21]

Regarding use cases consuming real-world databases developed using the vocabularies aforementioned, two different applications have been developed. namely *PerfectSchool*[22] and *Equipment*[23]. The former is a mobile application intended to provide useful information on schools in France, while the latter is a facet view by categories of facilities in France, specifically in the city of Toulouse.

---

17 `http://data.ign.fr/def/ignf`

18 A service to lookup CRS in RDF can be found at `http://www.eurecom.fr/~atemezin/ignf-lookup/`

19 `http://data.ign.fr/def/geofla#`

20 `http://data.ign.fr/def/topo`

21 `https://github.com/gatemezing/ign-iswc2014/blob/master/vocabularies/mappingsGeonames.ttl`

22 `semantics.eurecom.fr/datalift/PerfectSchool/`

23 `http://semantics.eurecom.fr/datalift/Equipment/`

4 Publishing administrative units (GeoFla)

As a dataset dedicated to administrative units, GEOFLA®is very likely to be reused by other datasets, either by reusing directly its URIs for georeferencing needs, or by reusing its description of administrative units - labels, properties and geometries - for interlinking purposes.

4.1 Data conversion

Geofla is delivered as a set of 4 shapefiles that describe the boundaries and properties of administrative units of mainland France (for CRS reasons, overseas territories are delivered within different shapefiles) : communes, cantons, arrondissements and departements. For the sake of our application, we have generated another shapefile describing regions by aggregating the geometries of the instances of departments based on their region's foreign key value. This dataset is updated every year. Publishing this data in RDF with unique identifiers on the Web will ease the interlinking with some existing datasets describing French boundaries in the wild. We follow a two steps conversion: we use the SHP2RDF module of Datalift to obtain a raw RDF from shapefiles, and the RDF2RDF module of Datalift using a set of SPARQL construct queries[24] for getting a refined RDF datasets using suitable vocabularies.

4.2 URI design policy

One of the requirements to publish data is to have unique ids and stable URIs[25] . Since our legacy databases have unique IDs to refer to the objects, we had to make sure they were unique at Web level. Thus, the base scheme for vocabularies URIs is: `http://data.ign.fr/def/`. Besides, the base schema for identifying a real world resource uses `http://{BASE}/id/`. For example, IGN main buildings are located in the commune with the URI `rgeofla:commune/94067`, corresponding to Saint-Mandé, and `rgeofla:departement/94` corresponds to the department "Val de Marne" to which the commune belongs.

4.3 Interlinking with existing GeoData

We interlinked our datasets with NUTS, DBpedia FR[26] and GADM datasets. SILK [6] is used to interlink the departments in our dataset with departments in DBpedia FR, using labels and INSEE Code. We obtained 93 matches (all correct) while three are missing for the departments 07, 09 and 75[27]. The LIMES tool[28] is then used to perform the rest of the interlinking tasks [7] with the trigrams function based on the labels with restriction to France.

24 `https://github.com/gatemezing/ign-iswc2014/tree/master/rdf2rdf`
25 `http://www.w3.org/TR/ld-bp/#HTTP-URIS`
26 `http://fr.dbpedia.org/`
27 `https://github.com/gatemezing/ign-iswc2014/tree/master/interlinking/matched`
28 `https://github.com/AKSW/LIMES.`

– Geofla-RDF with DBpedia FR: **23 252** links obtained. This results show the missing of nearly 13 435 communes not correctly typed in DBpedia FR as `Spatial Feature` or `Place`, or not having a French Wikipedia entry.
– Geofla-RDF with GADM (8 314 443 features): **70** links obtained: 10 communes, 51 departments and 9 regions. The property `gadm:in_country` is used to restrict the interlinking to France. E.g.: The city of Saint-Alban in Quebec is a commune in France.
– Geofla-RDF with NUTS (316 236 triples): Using a *"naive"* script with `trigrams` function on `geofla:Commune/rdfs:label` and `spatial:Feature/ramon:name` reveal two odd results located in Germany and Switzerland. The latter being the *JURA* and the former named *"Celle"*. In order to remove those odd effects, we add another restrictions based on ramon:code by filtering the ones located in France (136 features) . The final matchings give a total of **105** correct links: 14 communes, 75 departments and 16 regions.

The above results show good precision of the matching algorithm (score above 0.98) and a rather low recall value with DBPedia-FR (0.627). The few number of matched entities is likely due to the low coverage of French features in the datasets.

5 Publishing French Gazetteer

In this section, we present some first tests of converting BDTOPO® into RDF and interlinking with LinkedGeoData using LIMES. The results confirm the need for geographic publishers to publish georeference data on the Web.

Data conversion, URIs and Interlinking: Shapefiles are converted into RDF using the same two conversion process as for GEOFLA®. The URIs for each resource follow the pattern: `rtopo:CLASS/ID` for the feature, while `rtopo:geom/CLASS/ID` is used to reference the geometry of the resource. The gazetteer dataset in RDF is part of BD TOPO® database consisting of 1,137,543 triples (103,413 features). We chose LinkedGeoData (LGD) [29] to perform the alignments using the main class `lgdo:Amenity`[30] (5,543 001 triples), as they are closed to the features contained in the gazetteer. We perform the interlinking on the geometries using the `hausdorff` metric of LIMES tool. A total of **654** alignments was obtained above the threshold (0.9). This relatively low number of hits can be explained by the coverage of French data in LGD, and the subset of BDTOPO® used for the interlinking. Table 2 provides details of the alignments with subclasses of Amenity.

6 Opportunities and Challenges

The need for interoperable reference geographic data to share and combine georeferenced environmental spatial information is particularly acknowledged by

[29] http://linkedgeodata.org/sparql
[30] http://linkedgeodata.org/ontology/

LGD Class	#links matched
lgdo:Shop	252
lgdo:TourismThing	30
lgdo:Craft	3
lgdo:AerowayThing	37
lgdo:AerialwayThing	11
lgdo:EmergencyThing	56
lgdo:HistoricThing	257
lgdo:MilitaryThing	8

Table 2: Interlinking results using the Hausdorff metric of LIMES tool between LinkedGeoData and toponyms in the French Gazetteer

the INSPIRE Directive. For geographic data producers, the benefit of publishing their data on the Web according to Linked Data (LD) principles is twofold. On the one hand, their data are interoperable with other published datasets and they can be referenced by external resources and used as spatial reference data, which would not have been straightforward when published according to spatial data infrastructures (SDI) standards. On the other hand, the use of semantic Web technologies can help addressing interoperability issues which are not solved yet by geographic information standards. Moreover, there are different types of license policies to access data at IGN (e.g.: research purpose, commercial use, access on demand, etc.), with some of them not necessary "open" or free to access: (e.g. BD TOPO®). Although there is a clear understanding of the benefits of publishing and interconnecting data on the web, ongoing investigations on how to combine licenses on datasets are under consideration at IGN. Two solutions are under investigation: (i) different license policies attached to datasets and (ii) the use of a security access mechanism on top of the datasets granting access based on a predetermined configuration on named graphs and resources. According to Linked data principles URIs should remain stable, even if administrative units change or disappear. This implies adapting the data vocabulary in order to handle data versioning and real world evolutions. This issue will be addressed in a future work, as we plan to release a spatio-temporal dataset describing the evolution of communes since the French Revolution. Another issue deals with the automation of the whole publication process, from traditional geographic data to fully interconnected RDF data. The last issue deals with the use of multiple geometries for describing a geographic feature: geometries with different levels of detail, different CRS, different representation choices. This has been superficially addressed in our use case with the use of both polygons and points for representing respectively the surface and the centroid of departments, but should be further investigated for both query answering and map design purposes.

7 Conclusions

In this article, we proposed to contribute to the georeferencing of datasets published on the Web of data by providing two vocabularies designed for representing structured geometries defined with coordinates expressed in any CRS, as well as referencel geodata resources published under `data.ign.fr`, namely CRS's dataset, the French administrative units dataset and part of the French gazetteer dataset. So far, the French units are interconnected with the French statistical datasets, and reused in metadata fields used by the `www.datalocale.fr` portal for defining the geographic extent of each dataset[31].

Acknowledgments

This work has been partially supported by the French National Research Agency (ANR) within the Datalift Project, under grant number ANR-10-CORD-009.

References

1. A. G. Atemezing and R. Troncy. Comparing Vocabularies for Representing Geographical Features and Their Geometry. In *5th International Terra Cognita Workshop*, Boston, USA, 2012.
2. A. de León, L. M. Vilches, B. Villazón-Terrazas, F. Priyatna, and O. Corcho. Geographical linked data: a Spanish use case. In *International Conference on Semantic Systems (I-SEMANTICS'10)*, Graz, Austria, 2010.
3. J. Goodwin, C. Dolbear, and G. Hart. Geographical Linked Data: The Administrative Geography of Great Britain on the Semantic Web. *Transactions in GIS*, 12:19–30, 2008.
4. INSPIRE Thematic WG CRS and Geographical Grid Systems . Guidelines INSPIRE Specification on Coordinate Reference Systems , 2009. `http://inspire.ec.europa.eu/documents/Data_Specifications/INSPIRE_Specification_CRS_v3.0.pdf`.
5. International Organization for Standardization . ISO 19125-1, Geographic information- Simple feature access - Part 1: Common architecture, 2004.
6. A. Jentzsch, R. Isele, and C. Bizer. Silk-generating rdf links while publishing or consuming linked data. In *Poster at the International Semantic Web Conference (ISWC2010), Shanghai*, 2010.
7. A.-C. Ngonga Ngomo. Orchid - reduction-ratio-optimal computation of geo-spatial distances for link discovery. In *Proceedings of ISWC 2013*, 2013.
8. F. Scharffe, G. Atemezing, R. Troncy, F. Gandon, S. Villata, B. Bucher, F. Hamdi, L. Bihanic, G. Képéklian, F. Cotton, J. Euzenat, Z. Fan, P.-Y. Vandenbussche, and B. Vatant. Enabling linked-data publication with the datalift platform. In *26th Conference on Artificial Intelligence (AAAI-12)*, 2012.
9. P. Vandenbussche, B. Vatant, and L. Rozat. Linked open vocabularies: an initiative for the web of data. In *QetR Workshop, Chambery, France.*, 2011.

[31] `http://www.datalocale.fr/drupal7/dataset/ens-cg33`