The case for a standardised CRS ontology*

Timo Homburg1,4,†, Frans Knibbe2,†, Ghislain Atemezing3,†, Nathalie Abadie4,† and Luis Moreira de Sousa5,†

1Hochschule Mainz, University of Applied Sciences, Lucy-Hillebrand-Straße 2, 55128 Mainz
2Triply, Marconiweg 25, 1401 VG Bussum, The Netherlands
3ERA - European Union Agency for Railways, 120 Rue Marc Lefrancq, 59307 Valenciennes, France
4LASTIG, University Gustave Eiffel, IGN-ENSG, 73 Avenue de Paris, F-94165 Saint-Mandé, France
5ISRIC - World Soil Information, Building 101, Droevendaalsesteeg 3, 6708 PB Wageningen, The Netherlands

Abstract
This article presents the case for a standardised web ontology for Coordinate Reference Systems (CRS), a component currently missing in the semantic description of (geo-)spatial data on the web. Such an ontology should simplify the access and interpretation of CRSs and their parametric definitions. Motivations are laid out, presenting the needs of developers, data curators, and users of spatial data. Possible approaches to developing such an ontology and use cases that may facilitate tackling customized coordinate reference systems incorporating the ontology model are described.

Keywords
Coordinate Reference Systems, Geographic Information Systems, Spatial Reference Systems, Ontology, GeoSPARQL

1. Introduction
Spatial data have played a central role in the Web of Data since its inception, providing an intuitive way of linking datasets[1]. Space and time are elements present in almost all types of data [2]. Provided they can be interpreted unambiguously, these data can be brought together at large in the same context. To this end, Coordinate Reference Systems (CRS) are an essential component [3] since they provide the means to correctly interpret and process geometry coordinates. Over the centuries, countries and mapping agencies have created hundreds of different CRS types, varying by their area of validity, type (1D, 2D, 3D), associated cartographic projection, and various other parameters. Many of these CRSs remain in active use today. Dictionaries of place names, also known as gazetteers, have long been part of the Web of Data as an intuitive way of linking datasets[1].
Data [4]. The Geonames service\(^1\) and the Getty Thesaurus of Geographic Names\(^2\) are two well-known examples. However, these provide, at best, rough locations on the Earth. CRSs, on the other hand, provide high-precision locations and are of major interest in ensuring the consistency and interoperability of coordinates published on the Web. Although their importance is clearly highlighted by the GeoSPARQL specification [5, 6] of the Open Geospatial Consortium (OGC), that standard does not provide a suitable ontology for the provision and exchange of CRS definitions and parameters, merely recommending the use of URIs as identifiers.

This article surveys existing initiatives to publish CRS descriptions on the Web, their advantages and drawbacks (Section 2). The expected benefits of a standard web ontology for the description of CRSs are listed (Section 3) and several possible use cases are presented thereafter. The article closes with directions for future work (Section 4).

1.1. Definition

This section defines the essential terminology required to understand CRSs:

- **spatial reference system**: A spatial reference system (SRS) is a system for establishing spatial position. A spatial reference system can use geographic identifiers (place names, for example), coordinates (in which case it is a coordinate reference system), or identifiers with structured geometry (in which case it is a discrete global grid system).

- **coordinate system**: A coordinate system is a set of mathematical rules for specifying how coordinates are to be assigned to points.

- **datum**: A datum is a parameter, or set of parameters, defining the position of the origin, the scale, and the orientation of a coordinate system.

- **coordinate reference system**: A coordinate reference system (CRS) is a coordinate system that is related to an object by a datum.

- **CRS registry**: A CRS registry is a collection of descriptions of coordinate reference systems.

A CRS may provide information on:

- the unit of measure expressed by the coordinates;
- the spatial object on which the coordinates exist;
- capabilities in coping with 2D or 3D coordinates;
- the properties of an associated cartographic projection.

The OGC Reference Model [2] lists three types of SRSs to describe locations:

1. Geographic identifiers, like place names or postal codes;
2. Coordinate reference systems, which allow coordinates to be interpreted unambiguously;
3. Linear referencing, which allows locating things or phenomena along linear segments.

\(^1\)https://www.geonames.org/
\(^2\)https://www.getty.edu/research/tools/vocabularies/tgn/
1.2. Motivation

Even though CRSs are an essential part of any geometric data publication in RDF, a standardised vocabulary for their representation (with all relevant parameters) is still lacking in the context of Linked Open Data (LOD).

A decade ago, the GeoSPARQL 1.0 standard issued by the OGC created a formal and thorough framework to encode and publish geospatial data with RDF [5]. This standard offers semantically sound definitions of common geospatial concepts, such as Features, Collections, or Geometries. However, it highlights a single CRS as default usage in Well-Known Text Literals [7], the axes-swapped CRS84, whose definition is published by the OGC, but in GML format\(^3\). Because of this, CRS84 is also the most, and often the only implemented CRS in triple store implementations according to [8]. The expression of any other CRS is left entirely to the user, be it on the format or publication medium.

The need for a web ontology for the semantic publication of CRSs is apparent. It is a necessary milestone for the wider adoption of RDF for the publication of spatial data.

2. Related Work

Publishing spatial data on the LOD cloud requires users to be able to correctly interpret the coordinates that describe the shape and location of geometry. These coordinates can be expressed in various CRSs, depending on the area of the world covered by the data, their producer, and their intended use. Some coordinate systems are well suited to precision distance calculations, while others are preferable for producing statistical maps. For high-precision geography, a CRS free from the influence of plate tectonics is preferred. This section digests related work on semantic CRS representations and CRS ontology development.

2.1. OGC/ISO standard conceptual schema

The ISO 19111 standard[9] and the OGC abstract specification ”Referencing by coordinates”[10] provide a conceptual schema for the description of CRS parameters. The two standards are mutually consistent. They cover all kinds of CRS whose coordinate values do not change over time unless this change can be defined with monotonic parameters. Although the schema originates from the geography domain, it allows for the expression of engineering CRSs, which can be used for geometry that is not directly earth-related.

The conceptual schema also includes the operations that change coordinate values (like coordinates transformations from one given CRS to another) and some CRS metadata. The schema is expressed in UML and is used in standards such as GML [11] and the Well-known text representation of coordinate reference systems [12]. An official CRS web ontology would need to be fully semantically compatible with this schema in order to guarantee interoperability with many existing information systems.

\(^3\)https://www.opengis.net/def/crs/OGC/1.3/CRS84
2.2. CRS identifiers and registries

The INSPIRE CRS specification assigns identifiers to the coordinate systems that it recommends [13]. Its corresponding implementation specification [14] recommends implementing a registry for the dissemination of CRS identifiers and their associated descriptions. In the ISO TC-211 series of standards for geographic information, a registry is an “information system on which a register is maintained”; a register is defined as a "set of files containing identifiers assigned to items with descriptions of the associated items" [15]. The INSPIRE implementation specification, therefore, advises the URIs proposed by the OGC to be used as CRS identifiers.

2.2.1. URIs to identify coordinate reference systems on the Web of Data

To foster the adoption of URIs as CRS identifiers, the OGC proposes URIs to identify the most commonly used CRSs on the Web, including the WGS84 ensemble and the CRSs recommended by the INSPIRE Directive. These redirect to descriptions of the corresponding reference coordinate systems extracted from the EPSG geodetic parameters registry, compliant with the ISO 19111 standard [9]. This provides a conceptual model for the description of CRSs and the geodetic objects they are based on. Thus the URI http://www.opengis.net/def/crs/EPSG/0/4326 returns the GML [11] description of the WGS84 ensemble, as provided by the EPSG. However, this initiative does not cover all existing CRSs, and their descriptions provided are not encoded in RDF [16].

2.2.2. State-of-the-art of CRS registries

Several Web services provide access to much more comprehensive registries of CRSs:

- **EPSG Geodetic Parameter Registry**: is maintained by the Geomatics Committee of the International Association of Oil and Gas Producers. It allows queries on a dataset describing the geodetic parameters of several thousand CRSs. However, no direct access by URI dereferencing and content negotiation is possible.

- **EPSG.io**: provides access to the descriptions of the CRSs defined in the EPSG dataset using dereferenceable URIs, concatenating the service authority with the EPSG identifier. E.g. the URI http://epsg.io/4326 dereferences to an HTML page with the WGS84 description.

- **European Reference Coordinate System Service**: provides access to ISO 19111 compliant descriptions of the main European CRSs. Access by dereferencing URIs is not possible, and CRS descriptions are only available in HTML.

- **SpatialReference.org**: registry providing access to the description of many CRSs by dereferencing URIs. These URIs are built from the CRS identifiers on other registries

---

5 [https://epsg.org/](https://epsg.org/)
6 [https://www.iogp.org/our-committees/geomatics/](https://www.iogp.org/our-committees/geomatics/)
7 [https://epsg.io/](https://epsg.io/)
9 [https://spatialreference.org/](https://spatialreference.org/)

- **French national mapping agency (IGN France) registry**: consistent with the requirements of the INSPIRE Directive, IGN France publishes and maintains a registry of CRSs. These are identified by URIs composed of short names rather than numerical codes to designate geodetic resources, e.g., datums, ellipsoids, axes, meridians, etc. The URI https://registre.ign.fr/ign/IGNF/crs/IGNF/NTFLAMB2E dereferences to the Lambert zone II extended CRS.

None of the registries listed above provide CRS descriptions in the RDF model. Thus, interpreting CRS data is impossible using Linked Data tools and methods such as SPARQL or simple dereferencing of URIs describing CRS.

### 2.3. The IGN CRS ontology

In 2014, the French National Mapping Agency, known as the *Institut National de l’Information Géographique et Forestière (IGN)*, proposed a set of ontologies that extend the GeoSPARQL standard to publish its data on French administrative units on the Web of Data [17].

The first ontology\(^{11}\) was published to describe the administrative division of the territory: municipalities, cantons, arrondissements, *départements*, and regions. A second ontology\(^{12}\) has been proposed to represent the geometries associated with these administrative units in a structured way. It extends the main GeoSPARQL geometry classes to remain compatible with this standard. For example, the class `geom:Geometry` in this vocabulary is defined as an `owl:subClassOf sf:Geometry`\(^{13}\). As such, its concrete subclasses can be associated with WKT descriptions of geometries with the `geo:asWKT` property defined by GeoSPARQL\(^{14}\). But they can also detail the coordinates describing these geometries with dedicated properties so their values can be directly queried. For example, this vocabulary provides `geom:x`, `geom:y`, and `geom:z` properties to represent an instance of the class `geom:Point`, or `geom:points` to link a resource of type `geom:Linestring` to an ordered list of resources of type `geom:Point`.

Moreover, the `geom:Geometry` class is also defined as equivalent to an anonymous class populated by all resources associated with exactly one resource of type `ignf:CRS` by the `geom:crs` property. This `ignf:CRS` class is defined by a third ontology\(^{15}\), proposed to publish the registry of CRSs defined and maintained by IGN on the Web of Data. This ontology is based on the ISO 19111 standard [9], focusing on the types of CRSs managed by IGN France. As an example, this ontology defines the concept of `SingleCRS`\(^{16}\) but not that of `EngineeringCRS`\(^{17}\) as it is not useful to represent in IGN’s CRS registry. In addition, this ontology does not redefine the

\(^{10}\)https://registre.ign.fr/ign/IGNF/IGNF/

\(^{11}\)`geofla:http://data.ign.fr/def/geofla#`

\(^{12}\)`geom:http://data.ign.fr/def/geometrie#`

\(^{13}\)The prefix sf is used for `http://www.opengis.net/ont/sf#`

\(^{14}\)The prefix geo is used for `http://www.opengis.net/ont/geosparql#`

\(^{15}\)`ignf:http://data.ign.fr/def/ignf#`

\(^{16}\)A single CRS is a CRS consisting of one coordinate system and one datum.

\(^{17}\)An engineering CRS is a CRS based on an engineering datum, i.e. a local datum based on a local reference. It may be used to describe relative locations in a local CRS or coordinates in a CRS centred on a moving object.
concepts and properties for which well-known ontologies already exist. For example, the QUDT ontology\textsuperscript{18} is reused to describe the units of measurement. IGN’s CRS registry 2.1.3 has been transformed from GML format to RDF and represented according to this ontology. It has been published on the Web of Data and is accessible through the same SPARQL endpoint\textsuperscript{19} as IGN’s geodata about French administrative units. The URI http://data.ign.fr/id/ignf/crs/RGF93LAMB93 dereferences to the RDF representations of the Lambert 93 CRS, the legal projected coordinates reference system for mainland France. The query presented in Listing 1 retrieves the URIs and names of all the projected CRSs defined and maintained by IGN France.

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX ignf: <http://data.ign.fr/def/ignf#>

SELECT ?s ?l
WHERE {
  ?s rdf:type ignf:ProjectedCRS .
  ?s rdfs:label ?l .
  }
```

Listing 1: A SPARQL query on the IGN France’s RDF registry

### 2.4. ISO 19111 web ontologies

The ISO has made web ontologies available online, based on the ISO-19111 standard\textsuperscript{20}. These ontologies were automatically derived from the XML schemas of the standard. Therefore, they do not have the quality expected of a standard CRS web ontology. For instance, many URIs can not be resolved, many terms are undefined, many blank nodes have unclear meanings, and the good practice of using existing web ontologies is not followed. The ISO 19111 web ontologies seem to be the result of a one-off experiment. They do not seem fit for purpose and in their current state, would be a poor foundation for CRS definitions and CRS registries on the Semantic Web.

### 2.5. Proj4RDF

The proj4rdf project\textsuperscript{21} tackles a use case to extract CRS-related data from existing libraries such as PROJ\textsuperscript{22}. In addition, proj4rdf models vocabularies for a variety of projection types and aims to include multiple coordinate systems that go beyond those used in geospatial settings, e.g. interstellar coordinate systems. The results of the approach are:

- An RDF representation of the PROJ database, partially aligned with the ISO specification and rendered as HTML\textsuperscript{23}.
- A draft for a JSON-LD context based on the proposed vocabulary.
- The distribution of extension vocabularies\textsuperscript{24} for:

\textsuperscript{18}http://qudt.org/1.1/schema/qudt
\textsuperscript{19}The endpoint http://data.ign.fr/id/sparql can be queried through https://yasgui.triply.cc/
\textsuperscript{20}https://def.isotc211.org/ontologies/iso19111/
\textsuperscript{21}https://github.com/situx/proj4rdf
\textsuperscript{22}https://proj.org/
\textsuperscript{23}https://situx.github.io/proj4rdf/data/def/crs/EPSG/0/4328/index.html
\textsuperscript{24}https://situx.github.io/proj4rdf/
– Projections: Parameters for projection functions;
– Coordinate Systems;
– Grid types for DGGS and GeoCoding types;
– Planets: Specification for planet spheroids;
– SRS Applications: Concepts that describe the typical application cases of a CRS.

• A collection of SHACL shapes for validation.

Together with the IGN CRS ontology, these two projects could provide a good foundation for a standard CRS ontology. However, the conversion between the ontology model and other established ways of providing CRS data, such as WKT, is yet to be developed.

3. Benefits and use cases

Having a standard CRS ontology on the Web will have major benefits that empower many use cases. Here, we list four benefits and the use cases that each of them makes possible.

3.1. Provision of CRS semantics on the Web

A standard CRS ontology will provide an RDF/RDFS/OWL representation of all concepts related to coordinate reference systems. Various data and domain models for CRS definitions have been issued by authorities such as the OGC and ISO. Reference software packages, such as PROJ, feature a de facto standard data model. All of these are, at best, semantically defined in an electronic document. Web-based and dereferenceable semantic definitions of CRS concepts and parameters would make for a relevant advancement in the communication and correct use of CRSs.

Use cases

1. Provide human readable definitions of CRS elements directly from geometric data to facilitate understanding and avoid usage errors.
2. Provide a seamless link between geometric data and how their coordinates should be interpreted.
3. Enable reasoning on CRS elements.
4. Enable expression of custom CRSs.
5. CRS data will be usable by both people and machines/algorithms.
6. Allow the ISO-19111 model to be easily extended, for example, for extraterrestrial CRSs or other customized extensions.
7. Allow CRS specifications to be used in dataset metadata, optionally removing the need for specifying the CRS for individual geometries.
8. Allow all CRS elements to be used in (federated) SPARQL queries. For example, filter by unit of measurement or by applicable area.
9. Enable CRS recommendations based on the extent of the concerned spatial dataset and coordinate precision.
3.2. Enable publication of CRS registries on the Web

Once a standard CRS web ontology is brought online, expressing any CRS in RDF will be possible. In turn, this will enable the publication of collections of RDF-based CRS definitions in CRS registries, allowing data and datasets to use common URIs to reference CRSs.

Use cases

1. An official CRS registry by e.g. the OGC can be published, providing common URIs for common CRSs that can be resolved to RDF data.
2. Remove the need to replicate and update the parameters of common CRSs to data stores.
3. Well-known official IRIs can be used to match CRSs in web searches or federated searches.
   Example: find all datasets with a CRS that matches an interactive web map.
4. Official national grids can be published by national mapping and cadastral agencies.
5. Enable validation of coordinate data, e.g. via SHACL. For example: check if all coordinate values are within the extreme values.
6. Allow CRS specifications to be used in metadata standards, GeoDCAT-AP\textsuperscript{25} for example.
7. Stand-alone systems that do not publish data on the Web can benefit from access to up-to-date CRS data without needing local copies that run the risk of being outdated.
8. Allow provision of JSON-LD contexts for established JSON-based CRS schemes.

3.3. Complement GeoSPARQL

GeoSPARQL is arguably the most important standard for spatial data on the Web. It offers ways to work with geometry, which relies on CRS data, but the standard does not include CRS semantics. Therefore, a standard CRS ontology would be a welcome complement to GeoSPARQL.

Use cases

1. CRS registries can provide targets for a new property of the GeoSPARQL \texttt{Geometry} class that identifies the CRS.
2. GeoSPARQL currently has no way to encode geometry in RDF. It relies on non-RDF serialisations to express geometry. A standard CRS ontology would contain definitions of the \texttt{coordinate} and \texttt{coordinate reference system} concepts, which are two basic components of the definition of \texttt{Geometry}. The envisioned ontology would thus strengthen the definition of geometries as RDF resources.
3. (Federated) GeoSPARQL queries become feasible with geometries that use a custom CRS (a CRS not included in any CRS registry).

3.4. Increase interoperability of spatial data on the Web

Many types of spatial data, not only geographical data, use coordinates and therefore need CRS specifications. A standard CRS ontology can provide increased semantic and operational interoperability between all coordinate-based data.

Use cases

1. Geographic geometry and other types of geometry can use the same CRS semantics.
2. Facilitate georeferencing with local CRSs.
3. Make coordinate transformations possible with Linked Data tools.
4. CRS semantics can be made available to knowledge domains outside of geoinformatics, e.g. in the cultural heritage domain.
5. Historical coordinate reference systems can be published using the same semantics as modern CRSs. For example, the CRS parameters of the Verniquet map, a large-scale map of Paris produced on the eve of the French Revolution, could be published in RDF [19]. This would make the CRS available to the scientific community for geo-referencing with subsequent plans of Paris, which were based on the CRS created by Edme Verniquet for the purposes of surveying his map.

4. Conclusions and future work

This article argues for establishing a formalised and standardised ontology to represent CRSs. Existing approaches for developing a CRS ontology are presented, therefore collecting the necessary components that a standardised ontology model is expected to fulfil. Further, the article illustrates the usefulness of such ontology with a collection of different use cases, showing how the CRS model could be integrated into a JSON-LD context and with already existing standards such as GeoSPARQL.

Future work would need to see the formation of a standardisation effort, possibly within the OGC. This standardisation effort should be compatible with already existing non-RDF standardisation efforts, and incorporate all related work mentioned in this document. Legacy models in UML would benefit from recent works on automatic serialisation into a formal OWL ontology, and a SHACL shape conforming to the SEMIC Style Guide [20]. In that context, proof-of-concept implementations will be needed to show the conversion between different CRS formats (RDF and legacy conceptual models) and an actual implementation that would show the application of the CRS ontology in practice. This could, for example, be the extension of a SPARQL query processing library with CRS RDF processing capabilities.

References


