Towards the FAIRification of Meteorological Data: a Meteorological Semantic Model

Amina Annane, Mouna Kamel, Cassia Trojahn, Nathalie Aussenac-Gilles, Catherine Comparot, Christophe Baehr

Towards the FAIRification of Meteorological Data: a Meteorological Semantic Model

Amina Annane[1 [0000−0002−7223−8478]], Mouna Kamel[1 [0000−0002−8739−0747]], Cassia Trojahn[1 [0000−0003−2840−005X]], Nathalie Aussenac-Gilles[1 [0000−0003−3653−3223]], Catherine Comparot[1 [0000−0002−0096−2766]], and Christophe Baehr[2]

[1] Toulouse University, France
firstname.lastname@irit.fr
[2] Météo-France, France
christophe.baehr@meteo.fr

Abstract. Meteorological institutions produce a valuable amount of data as a direct or side product of their activities, which can be potentially explored in diverse applications. However, making this data fully reusable requires considerable efforts in order to guarantee compliance to the FAIR principles. While most efforts in data FAIRification are limited to describing data with semantic metadata, such a description is not enough to fully address interoperability and reusability. We tackle this weakness by proposing a rich ontological model to represent both metadata and data schema of meteorological data. We apply the proposed model on a largely used meteorological dataset, the "SYNOP" dataset of Météo-France and show how the proposed model improves FAIRness.

Keywords: Meteorological data · FAIR principles · Semantic metadata

1 Introduction

Meteorology data is essential in many applications, including forecasts, climate change, environmental studies, agriculture, health and risk management. Their production is based on mathematical models that assimilate different data from several sources including weather stations, satellites and weather radars. While this data has been made available as open data through different portals, such as governmental portals (e.g., MeteoFrance[3], worldweather[4]), or associative or private portals (e.g., infoclimat[5], meteociel[6]), under open licenses, its exploitation is rather limited: it is described and presented with properties that are relevant for meteorology domain experts (data producers) but that are not properly understood and reusable by other scientific communities. For the latter, one of

[3] `https://donneespubliques.meteofrance.fr/`
[4] `http://worldweather.wmo.int/fr/home.html`
[5] `https://www.infoclimat.fr`
[6] `http://www.meteociel.fr`

the challenges is to find relevant data among the increasingly large amount of continuously generated data, by moving from the point of view of data producers to the point of view of users and usages. One way to reach this goal is to guarantee compliance of data to the FAIR principles (Findability, Accessibility, Interoperability, and Reusability) [21]. These principles correspond to a set of 15 recommendations to be followed in order to make data FAIR (Figure 1). A key step toward improving FAIRness of data is the use of semantic models (i.e., ontologies) for data and metadata representation [9].

While most efforts in data FAIRification are limited to describing data with semantic metadata, such a description is not enough to fully address all FAIR principles [12], in particular for promoting reuse of this data by other scientific communities. We propose to overcome this weakness through a richer representation of the meaning of meteorological data in a formal model which allows for sharing the semantic meaning with third-parties [13]. Contrary to existing works involving ontology population [14, 19, 2, 16, 1], and due to the characteristics of meteorological data, we do not transform all data into RDF but rather represent in a fine-grained way the data schema and its distribution structure. The contributions of this paper are as in the following:

- Proposing a semantic model representing both metadata and data schema of meteorological observation data.
- Combining existing vocabularies that follow themselves the FAIR principles.
- Instantiating the proposed model with a real meteorological dataset provided by Météo-France – the official weather service in France.
- Evaluating the FAIRness degree of this dataset without and with the proposed model showing how the proposed model improves the FAIRness degree.

The paper is organized as follows. §2 introduces the meteorological data specificities and presents the proposed model. §3 shows the instantiation and §4 presents the evaluation. §5 discusses the related work and §6 ends the paper.

2 Meteorological Semantic Model

In order to develop our model, we have followed the NeOn methodology scenario 3 "*Reusing ontological resources*" [20]: those cases in which ontology developers have at their disposal ontological resources useful for their problem and that can be reused in the ontology development. The methodology includes activities related to the reuse of ontologies (ontology search, ontology assessment, ontology comparison, ontology selection and ontology integration) in addition to the main activities (specification, conceptualization, formalization, implementation). For the sake of space, we focus here on the presentation of the model, without describing in detail the different activities of the methodology. In the following, we summarize the result of specification and knowledge acquisition in §2.1, and the result of ontology selection and integration in §2.2.

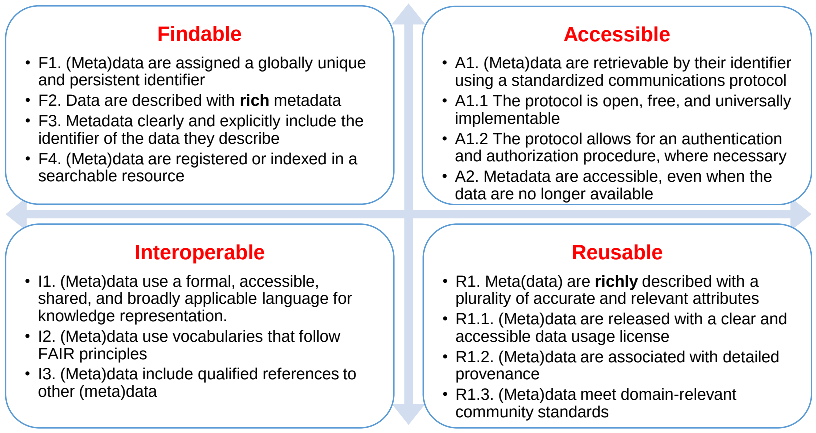

Fig. 1. FAIR principles [21].

2.1 Meteorological data characteristics

There exist different types of meteorological data: satellite data, model data that are computed using statistical models such as weather forecast data, radar data, etc. We focus on observation data referred to as "in situ" data. They are direct measurements of various parameters (temperature, wind, humidity, etc.) taken by instruments on the ground or at altitude from predefined locations:

– *Geospatial data:* the measure values must be localised, otherwise it is not fully exploitable. The localisation is usually defined using geospatial coordinates (latitude, longitude and altitude). The interpretation of these coordinates depends on the used coordinate reference system, hence the Coordinate Reference System (CRS) has also to be indicated.
– *Temporal data.* Each measurement is made at a specific time that must be noted with the measurement result (i.e., value). As for the geospatial, the temporal localisation is essential to the right interpretation of measurements.
– *Large volume data:* meteorological data are produced continuously. Within each station, several sensors are installed (thermometer, barometer, etc.). Each sensor generates multiple measure values with a frequency that differs from one measure to another (hourly, trihoral, daily, etc.).
– *Conform to WMO guidelines:* the measurement procedures, the types of sensors to use, the quality standards, and many other specifications are defined by the World Meteorological Organization (WMO). The latter provides detailed guides, such as *the guide to meteorological instruments and observation methods*[7] where there is a chapter for each measure describing all details about it. Nevertheless, to our knowledge, no semantic version of these guides exists.
– *Tabular data:* observation data are usually published in tabular format where measure values are organized according to spatio-temporal dimension. Ac-

[7] https://library.wmo.int/doc_num.php?explnum_id=10179

cording to a recent study made by Google, the tabular format is the most widespread format for publishing data on the Web (37% of the datasets indexed by Google are in CSV or XLS) [3].

2.2 Metadata and data representation

The proposed semantic model (Figure 2) represents both metadata and data schema of meteorological observation data, as described in the following. It has been implemented in OWL[8] and its consistency has been verified with the different reasoners implemented in Protégé (Hermit, ELK, and Pellet).

Metadata representation Making data FAIR requires first and foremost the generation of semantic metadata. Indeed, 12 out of the 15 FAIR principles refer to metadata as described in [21] (Figure 1). This metadata must remain accessible even if the data itself is no longer accessible. These 12 principles provide guidance on the categories of metadata: (i) descriptive metadata for data indexing and discovery (title, keywords, etc.); (ii) metadata about data provenance; (iii) metadata about access rights and usage licenses. Particularly for publishing data on the web, W3C recommends three other categories of metadata: (i) version history; (ii) quality; (iii) structure. Our goal is therefore to propose metadata model that covers these different categories, thus ensuring adherence to the principle on rich metadata. For metadata representation, our proposition relies on GeoDcat-AP, except structural metadata that are covered by CSVW.

GeoDcat-AP. It is a specification of the DCAT-AP vocabulary which is a specification of the W3C DCAT (Data CATalog vocabulary) recommendation. The choice of GeoDCAT-AP is motivated by the richness of the vocabulary metadata. In addition to allowing to describe data catalogs, it offers specific properties required to correctly interpret spatial data such as the geographical area covered by the data (`dct:spatial`), the reference coordinate system used (`dct:conformsTo`) to be chosen from a list defined by the OGC[9], as well as the spatial resolution (`dcat:spatialResolutionInMeters`) of the data. GeoDCAT-AP is also recommended by W3C/OGC to describe geospatial data on the Web [4].

CSVW. As pointed out in [12], it is essential for data reuse to represent the internal structure of the data. Since observation data are mostly tabular data, CSVW[10] is a suitable vocabulary. It resulted from the work of the W3C group on publishing tabular data on the web. It allows to define the different columns `csvw:Column` of a given `csvw:Table` (i.e., csv file) via the `csvw:Schema` concept. Moreover, it represents the interdependence between two tables. Indeed, it allows to represent if a column (or a set of columns) in a given csv file is a foreign key `csvw:ForeignKey` that references a column (or columns) of another CSV file.

[8] `https://www.irit.fr/recherches/MELODI/ontologies/DMO/core` `https://www.w3.org/IRIT/MELODI/Semantics4FAIR/DMO/core/`

[9] `http://www.opengis.net/def/crs/EPSG/`

[10] `https://www.w3.org/ns/csvw`

Web Annotation ontology (oa). It is a W3C recommendation for representing data annotations. As discussed in §2.1, all WMO members (i.e., states) use the same guides developed by WMO. It is used here to annotate parts of these guides that describe relevant metadata about measures. We use mainly three classes: (i) `oa:Annotation`, (ii) `oa:SpecificResource` to represent the annotated document, and (iii) the document-part annotated via the class `oa:RangeSelector`.

Data representation We have made the choice to not transform all data into RDF because it is: (i) expensive: transforming the data archived for decades requires human and physical resources, and (ii) not effective: it would result in a huge RDF graph that would not be effective for querying and accessing the data [11]. We rather represent the data schema and the RDF data cube (qb) vocabulary has been considered for that. In addition to qb, we have reused several domain and cross-domains ontologies for making explicit the semantics of measures and dimensions using concepts from the meteorological domain. It is worth to mention that we use CSVW to represent the syntactical structure of a tabular dataset distribution, while we use RDF data cube and domain ontologies to represent the semantics of the dataset independently of any data format.

RDF data cube (qb). It is a W3C vocabulary [4][11] dedicated to the representation of multi-dimensional data. qb is suitable in our case since observation data is multi-dimensional data organized according to spatio-temporal dimensions. It describes the multidimensional data schema using three subclasses of `qb:ComponentProperty`: (i) measures `qb:MeasureProperty`, (ii) dimensions `qb:DimensionProperty` according to which the measures are organized, and (iii) attributes to represent metadata `qb:AttributeProperty`. The `qb:concept` property allows to link a `qb:ComponentProperty` (i.e., measure, dimension or attribute) to the corresponding concept to explicit its semantics. We use this property to associate component properties to domain ontologies.

Domain ontologies. Meteorological data refers to concepts from the meteorological domain such as atmospheric parameters (e.g., temperature, wind speed), sensors (e.g. thermometer, barometer), etc. For expliciting their semantics, the following domain and cross-domain ontologies are used: SOSA [10] (Sensor, Observation, Sample, and Actuator), the reference ontology to represent observations; ENVO (Environment ontology) [5] and SWEET (Semantic Web Earth and Environment Technology ontology)[17], to designate the atmospheric parameters represented with qb as measurements; aws (ontology representing weather sensors)[12], to designate the type of weather sensor used to measure each atmospheric parameter; and QUDT, to specify the unit of measurement of each measurement.

2.3 New entities

In order to be able to fully address the ontology requirements, we have introduced the following entities:

[11] `https://www.w3.org/TR/eo-qb/`

[12] `https://www.w3.org/2005/Incubator/ssn/ssnx/meteo/aws`

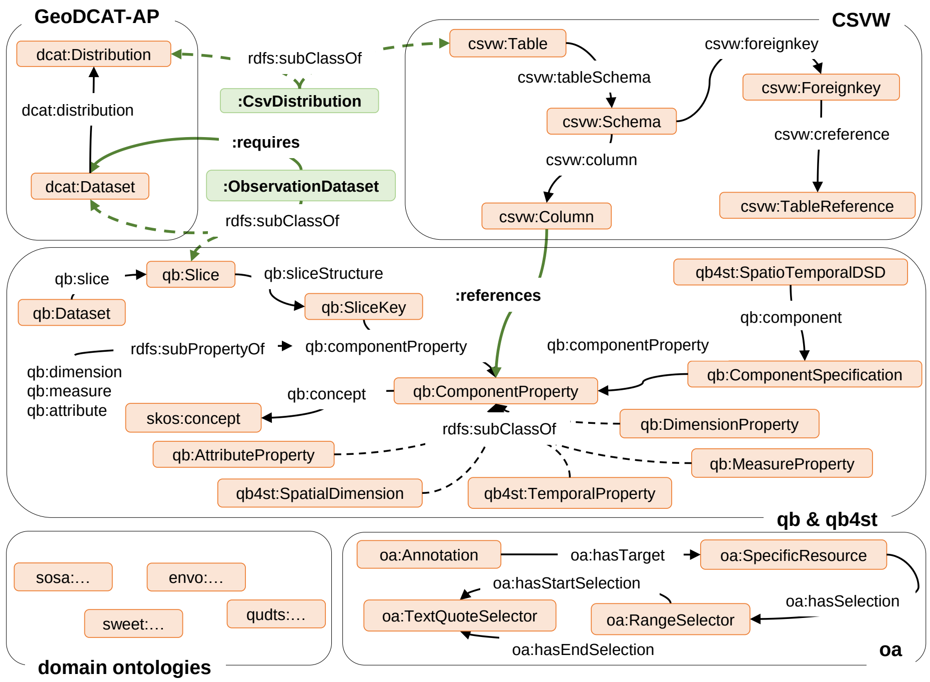

Fig. 2. Proposed modular ontology.

- *":ObservationDataset" concept*: given the continuous production of observation data, it is usually archived as fragments, where each fragment contains data of a given period (day, month, etc.). qb allows to represent the notion of a dataset (`qb:Dataset`) with multiples fragments (`qb:Slice`). However, GeoDCAT-AP offers only the possibility to represent a given dataset as a whole (`dcat:Dataset`). Since it is the same dataset fragment that we describe with GeoDCAT-AP and RDF data cube, we have defined a new concept `:ObservationDataset` that is a subclass of both `qb:Slice` and `dcat:Dataset`.
- *":CsvDistribution" concept*: it represents distributions in csv format. It is a subclass of both `csvw:Table` and `dcat:Distribution`.
- *":requires" property*: when a dataset X requires another dataset Y to be exploited, it is essential to enable the reuse of X to represent this dependency relation. As GeoDCAT-AP does not offer the possibility of representing such a relation between two datasets, we have added this new property.
- *":references" property*: it allows to associate each `csvw:Column` to an qb component property (i.e., `qb:MeasureProperty` or a `qb:DimensionProperty`): Thus, we explicit the link between the structural components (i.e., columns) and the data schema components (i.e., measures and dimensions). Moreover, data schema components are associated to domain ontology concepts, thereby explicit the semantic of each column too.

Table 1. Extract from SYNOP data.

numer_sta	date	pmer	ff	t	...
7005	20200201000000	100710	3.200000	285.450000	...
7015	20200201000000	100710	7.700000	284.950000	...
7020	20200201000000	100630	8.400000	284.150000	...
...	...	...	...	...	...

3 Use case: SYNOP dataset from Météo-France

3.1 Overview of SYNOP dataset

The SYNOP dataset is an open meteorological datasets provided by Météo-France on its data portal. It includes observation data from international surface observation messages circulating on the Global Telecommunication System (GTS) of the World Meteorological Organization (WMO). The choice of this dataset is motivated by the fact that this data is open and free, and it concerns several atmospheric parameters measured (temperature, humidity, wind direction and force, atmospheric pressure, precipitation height, etc.). These parameters are important for many scientific studies. The dataset[13] is described by seven items: (i) *description*: natural language summary that describes the content of the dataset, (ii) *conditions of access*: Etalab license[14] for the data, (iii) *means of access*: specifies that the data can be accessed via direct download, (iv) *download*: possibility offered to the user to download the data in csv format for a given date, (v) *download archived data*: similar to the previous item, but for a given month, (vi) *station information*: list of stations (station id, name) accompanied by a map displaying the location of these stations, and (vii) *documentation*, links to a data dictionary, to CSV and JSON files listing the meteorological stations of Météo-France (id_station, name, latitude, longitude, altitude). Table 1 shows an extract of the SYNOP data. The file contains 59 columns, the first two correspond to the station number and the date of the measurements made, the other 57 columns to the meteorological measurements.

3.2 Model instanciation

The SYNOP data archive consists of a set of monthly files since January 1996, where each file covers only the observations made in one month. The data of each monthly file corresponds to an instance of the `:ObservationDataset`. `:synop_dataset_feb20` is the instance corresponding to the SYNOP dataset of February 2020 (Table 1). For sake of space, we present a fragment of the instantiated model that show both metadata and data representation. It corresponds to the representation of a distribution of the dataset: `:synop_distribution_feb20`

[13] `https://donneespubliques.meteofrance.fr/?fond=produit&id_produit=90&id_rubrique=32`

[14] `https://www.etalab.gouv.fr/wp-content/uploads/2014/05/Licence_Ouverte.pdf`

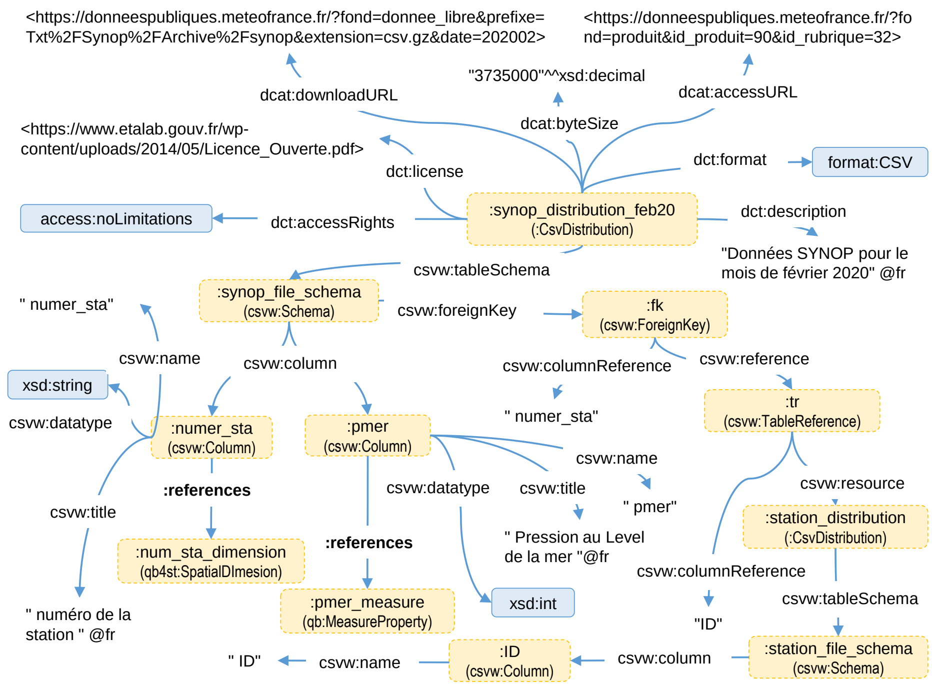

Fig. 3. Representing Synop_20 distribution using GeoDCAT-AP and CSVW.

(`:CsvDistribution`), as shown in Fig 3. This distribution is accessible and downloadable via URLs specified with properties `dcat:accessURL` and `dcat:downloadURL`; it is subject to an open license, the value of the `dct:license` property. Finally, the columns (e.g., `numer_sta` and `pmer`) of this file are characterized by their name (`csvw:name`), their label (`csvw:title`), their type (`csvw:datatype`) from the schema `:synop_file_schema` (`csvw:tableSchema`), etc. Note the representation of the foreign key `:fk` which connects the column "numer_sta" of the SYNOP data, to the column "ID" of the station data (`:station_distribution`) by passing through the instance `:tr` of `csvw:TableReference`.

In order to express the fact that all the monthly data are part of the same SYNOP dataset, we represent it as an instance of `qb:Dataset` (Figure 4). We have defined one spatial dimension `:station_dimension` and three temporal dimensions: `:year_dimension`, `:month_dimension`, and `:date_dimension`. The spatial or temporal nature of a dimension is specified using the concepts of the qb4st vocabulary `qb4st:SpatialDimension` and `qb4st:TemporalProperty`, respectively. Note that the year and month dimensions do not refer to existing columns, but are included in the date column. We have added them to instantiate `qb:Slice`. Indeed, the instantiation of a `qb:Slice` requires the definition of dimensions with fixed values, which are specified using the `qb:SliceKey` concept. In our case, the fixed dimensions for a monthly dataset are the year and the month, which have the values `month:FEB` and `year:2020` for the instance

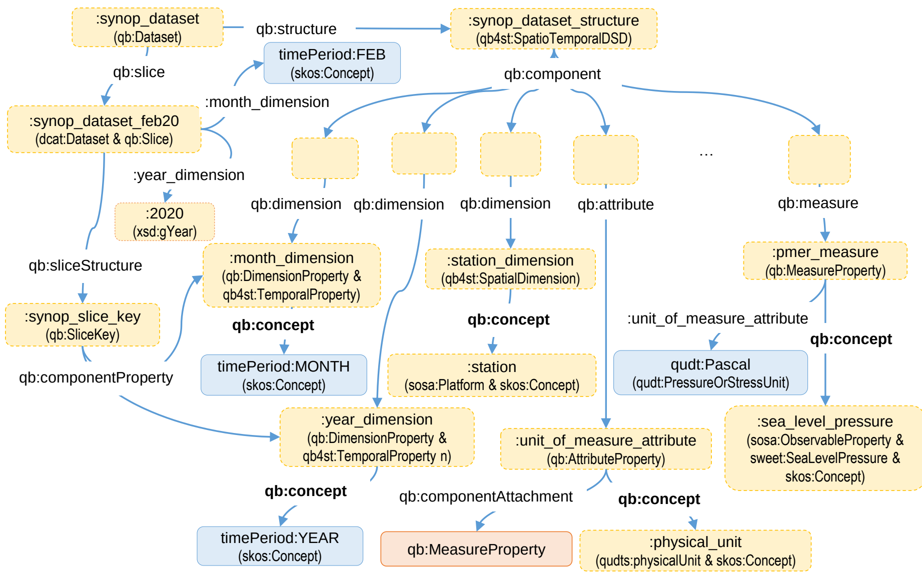

Fig. 4. Representing Synop_feb20 data using RDF data cube and domain ontologies.

`:synop_dataset_feb20`. In addition, although the station dimension is not directly a geographic coordinate, it is defined as an instance of `qb4st:spatialDimension` because it provides access to geospatial coordinates contained in the station file. Each dimension or measure is associated with a concept from domain ontologies via the `qb:concept` property. Thus, the measure `:pmer_measure` (the only one represented here while all 57 measures have been instantiated) is associated to the concepts `sosa:observableProperty` and `sweet:SeaLevelPressure` to explicit its meaning. Similarly, the measure `t` (see Table 1) is associated to the concept `ENVO:ENVO_09200001` which represents the `air temperature`. We have defined two attributes attached to `qb:Measure`: (i) `:unit_of_measure_attribute` associated to `qudts:physicalUnit` to represent the unit of measurement of each `qb:Measure`. This makes it possible to specify that the unit of measurement of `pmer_measure` is `qudt:Pascal`; (ii) :method_of_measure_attribute associated to `sosa procedure` to represent the procedure of measurement of each measure according to WMO guides that are digital books.

4 Evaluation

To evaluate the degree of FAIRness, we have chosen the framework *FAIR data maturity model* (FAIR data maturity model) proposed by the RDA [18]. For sake of space, we briefly present the evaluation of the SYNOP dataset using the FAIR maturity model using the 41 indicators that measure the level of a digital resource according to a FAIR principle, and the Fail/Pass evaluation method. Each indicator has a predefined priority: essential, important or recommended. The indicators were applied first considering the original description of the dataset,

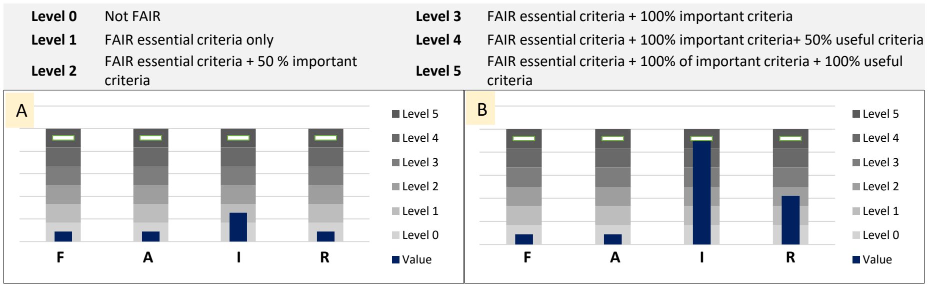

Fig. 5. Synop data evaluation: (A) without and (B) with semantic metadata.

and then considering the instantiation of the proposed model. The reader car refer to Zenodo[15] for a detailed evaluation report.

The first evaluation of the SYNOP dataset consisted in evaluating its original description (without the semantic model). This evaluation resulted in : i) level 0 for principles “F”, “A” and “R”, because at least one essential indicator was not satisfied for each of them; ii) level 1 for principle “I”, because no indicator is essential for this principle (Figure 5 (A)). As it stands, the SYNOP dataset is not FAIR. The data has been **re-evaluated** after generating the semantic metadata that describes it. This metadata significantly improves the FAIRness level, especially for the “I” and “R” principles (Figure 5 (B)). Although the re-evaluation of the “F” principle did not show any improvement, the model allows representing “rich” indexing metadata that satisfy “F2” principle. However, improving the “F” and “A” degree requires satisfying essential indicators that are beyond the abilities of any semantic model e.g., the generation of persistent and unique identifiers (“F1”), persistent metadata (“A2”), publishing metadata on searchable resources (“F4”), etc.

5 Related work

Hereafter, we present a summary on works related to the main subjects addressed in this work, focusing on works related to geospatial data.

Metadata representation. The importance of sharing geospatial data and describing them with rich metadata has been recognized for decades. Indeed, in 1999, Kim has published a comparative analysis of eight already existing metadata schemes for geospatial data. The INSPIRE (2007) directive defined a metadata schema, mainly based on the previous standards for describing the European geospatial data on web portals. Later, with the emergence of semantic web, semantic vocabularies have been developed to describe dataset metadata such as Dublin core, VoID, schema.org and DCAT. DCAT-AP was designed to ensure interoperability between European data portals. GeoDCAT-AP was initially developed to enable interoperability between geospatial data portals

[15] `doi.org/10.5281/zenodo.4679704`

implementing the INSPIRE directive, and those implementing DCAT-AP, by developing a set of mappings between the metadata schemes. In December 2020, a new version of GeoDCAT-AP was released, making this vocabulary a full-fledged specification for describing geospatial data catalogs on the Web[16].

Data representation. Several works have focused on the semantic representation of geospatial data [14, 19, 2, 16, 1]. The proposed models are generally a combination of a set of reference ontologies. In [14] the authors combined qb and SOSA to represent 100 years of temperature data in RDF. Similarly, in [19], the ontologies SOSA, GeoSPARQL, LOCN and QUDTS have been reused to represent a meteorological dataset with several measures (temperature, wind speed, etc.). In our case, given the characteristics of the meteorological data (§2.1), we did not transform Météo-France data into RDF. Representing all the data in RDF generates a huge graph which is not effective for querying the data [11]. Moreover, such a choice would require Météo-France to convert all its archives (some of them date back to 1872), which can turn out to be very expensive. Close to ours, [13] propose the Semantic Government Vocabulary, based on the different ontological types of terms occurring in the Open Government Data. They show how the vocabularies can be used to annotate Open Government Data on different levels of detail to improve "data discoverability".

FAIR principles and FAIRness evaluation. As discussed in [15, 8], semantic web technologies are the most in compliance to the implementation of FAIR principles. Since the appearance of FAIR principles in 2016, several frameworks have been proposed to evaluate the FAIRness degree of a given digital object. In most cases, the evaluation is performed by answering a set of questions – also called metrics or indicators in some works – or fill in a checklist[17] such as the "FAIR Data Maturity Model"[18] or "FAIRshake" [6]. Other works have proposed automated approaches for FAIRness evaluation [22, 7] based on small web applications that test digital resources against some predefined metrics.

6 Conclusion

This paper has presented an ontological model to represent both metadata and data schema of observational meteorological datasets. We have shown how the proposed model improves the adherence to FAIR principles. This work is part of an approach that aims to make meteorological data FAIR in general, and that of Météo-France in particular. The next step is to study the specifics of the data from the statistical models to enrich the current model if necessary. We plan as well to use the final model to generate the metadata and index it in data portals.

References

1. Arenas, H., Trojahn, C., Comparot, C., Aussenac-Gilles, N.: Un modèle pour l'intégration spatiale et temporelle de données géolocalisées. Revue Int. de Géomatique **28**(2), 243 (2018)

[16] `https://semiceu.github.io/GeoDCAT-AP/releases/2.0.0/`

[17] `https://fairassist.org/#!/`

2. Atemezing, G., Corcho, O., Garijo, D., Mora, J., Poveda-Villalón, M., Rozas, P., Vila-Suero, D., Villazón-Terrazas, B.: Transforming meteorological data into linked data. Semantic Web **4**(3), 285–290 (2013)
3. Benjelloun, O., Chen, S., Noy, N.F.: Google dataset search by the numbers. In: Proceedings of the 19th International Semantic Web Conference. pp. 667–682 (2020)
4. Linda van den Brink, e.: Best practices for publishing, retrieving, and using spatial data on the web. Semantic Web **10**(1), 95–114 (2019)
5. Buttigieg, P.L., Morrison, N., Smith, B., *et al.*: The environment ontology: contextualising biological and biomedical entities. J. Biomed. Semant. **4**, 43 (2013)
6. Clarke, D., *et al.*: Fairshake: Toolkit to evaluate the fairness of research digital resources. Cell Systems **9**(5), 417–421 (2019)
7. Devaraju, A., Huber, R., Mokrane, M., *et al.*: Fairsfair data object assessment metrics (Oct 2020), `https://doi.org/10.5281/zenodo.4081213`
8. *et al.*, A.J.: FAIR principles: Interpretations and implementation considerations. Data Intelligence **2**(1-2), 10–29 (2020)
9. Guizzardi, G.: Ontology, Ontologies and the "I" of FAIR. Data Int. **2**(1-2), 181–191 (2020)
10. Janowicz, K., Haller, A., Cox, S.J., Le Phuoc, D., Lefrançois, M.: Sosa: A lightweight ontology for sensors, observations, samples, and actuators. Journal of Web Semantics **56**, 1–10 (2019)
11. Karim, F., Vidal, M., Auer, S.: Compact representations for efficient storage of semantic sensor data. CoRR **abs/2011.09748** (2020)
12. Koesten, L., Simperl, E., Blount, T., Kacprzak, E., Tennison, J.: Everything you always wanted to know about a dataset: Studies in data summarisation. Int. J. Hum. Comput. Stud. **135** (2020)
13. Kremen, P., Necaský, M.: Improving discoverability of open government data with rich metadata descriptions using semantic government vocabulary. J. Web Semant. **55**, 1–20 (2019). https://doi.org/10.1016/j.websem.2018.12.009
14. Lefort, L., Bobruk, J., Haller, A., Taylor, K., Woolf, A.: A linked sensor data cube for a 100 year homogenised daily temperature dataset. In: Proceedings of the 5th International Workshop on Semantic Sensor Networks. vol. 904, pp. 1–16 (2012)
15. Mons, B., Neylon, C., Velterop, J., *et. al.*: Cloudy, increasingly fair; revisiting the FAIR data guiding principles for the european open science cloud. Inf. Serv. Use **37**(1), 49–56 (2017)
16. Patroumpas, K., Skoutas, D., Mandilaras, G.M., Giannopoulos, G., Athanasiou, S.: Exposing points of interest as linked geospatial data. In: Proceedings of the 16th International Symposium on Spatial and Temporal Databases. pp. 21–30 (2019)
17. Raskin, R.: Development of ontologies for earth system science. In: Geoinformatics: Data to Knowledge. Geological Society of America (01 2006)
18. RDA, F.D.M.M.W.G.: FAIR Data Maturity Model. Specification and Guidelines (Jun 2020). https://doi.org/10.15497/rda00050
19. Roussey, C., Bernard, S., André, G., Boffety, D.: Weather data publication on the LOD using SOSA /SSN ontology. Semantic Web **11**(4), 581–591 (2020)
20. Suárez-Figueroa, M.C., Gómez-Pérez, A., Fernández-López, M.: The neon methodology framework: A scenario-based methodology for ontology development. Appl. Ontology **10**(2), 107–145 (2015). https://doi.org/10.3233/AO-150145
21. Wilkinson, M., Dumontier, M., Aalbersberg, e.: The FAIR Guiding Principles for scientific data management and stewardship. Scientific data **3**(1), 1–9 (2016)
22. Wilkinson, M., Dumontier, M., Sansone, e.: Evaluating FAIR maturity through a scalable, automated, community-governed framework. Sc. Data **6**(1), 1–12 (2019)