Statistical Challenges Towards a Semantic Model for Precision Agriculture and Precision Livestock Farming

Dimitris Zeginis[1,2], Evangelos Kalampokis[1,2], and Konstantinos Tarabanis[1,2]

[1] Centre for Research & Technology Hellas, Information Technologies Institute, 6th km Charilaou - Thermi, Thessaloniki 57001, Greece
[2] University of Macedonia, Information Systems Lab, Egnatia 156, Thessaloniki, 54006, Greece
{zeginis, ekal, kat}@uom.edu.gr

Abstract. At the domains of agriculture and livestock farming big data come from diverse heterogeneous sources including structured data e.g. sensor data, weather/climate data and unstructured data e.g. drone/satellite imagery and maps. Big agricultural data can be used to provide predictive insights in farming operations, drive real-time operational decisions, and redesign business processes. However, the exploitation and integration of big agricultural data is not straightforward because: i) raw data (e.g. sensor data, satelite images) need to be further processed in order to extract valuable indicators (e.g. Normalized Difference Vegetation Index) or to be aggregated to the proper granularity level and ii) meta-data are needed to facilitate the exploration and integration of data (e.g. integrate data that have the same spatial and temporal coverage). In this paper, we study the characteristics of big agricultural data and propose a semantic model approach that facilitates their exploitation and integration. Towards this direction we study the semantic challenges that come up (e.g. granularity of data, data integration) and their potential solutions.

Keywords: Aggregated data, Linked data, big agricultural data, statistical Challenges, Precision Agriculture, Precision Livestock Farming

1   Introduction and motivation

Precision Agriculture (PA) uses intensive data collection and processing in time and space to make more efficient use of farm inputs, leading to improved crop production and environmental quality [5]. Similarly, Precision Livestock Farming (PLF) aims to create a management system based on continuous automatic real-time monitoring and control of production/reproduction, animal health and welfare, and the environmental impact of livestock production [2].

At both PA and PLF massive amounts of heterogeneous data are collected through numerous sources. For example, sensor data to measure soil electrical conductivity, satellite/drone images to see the state of crops at different parts of a field, weather/climate data from proprietary weather station or from meteorological institutes, and videos monitoring animal behaviour. These data can be used to provide predictive insights in farming and livestock operations, make predictions, drive real-time operational decisions, and redesign business processes.

A common approach to address the challenge of discovering data across numerous heterogeneous sources is to semantically annotate and publish their metadata. This could enable users to search based on a standardised approach in a transparent way for them. Towards this direction, standard vocabularies have been proposed such as the Data Catalog Vocabulary (DCAT).

In the case of PA and PLF, however, the majority of the raw data is of fine granularity coming from sensors, satellites, or drones. This type of data usually needs to be further processed to produce metrics that will be used in data-driven decision making scenarios. A typical index that is widely used in PA is the Normalized Difference Vegetation Index (NDVI), which quantifies vegetation. NDVI can be calculated from multispectral satellite images containing red and infrared channels. These new metrics could be of coarser granularity than the raw data depending on the problem at hand.

In some other cases, the raw data that are collected through various observation and sensing activities need to be aggregated before used in data-driven decision making scenarios. For example, raw meteo sensor data need to be aggregated in order to extract a valuable metric e.g. average temperature per day.

These data that are created from massive amounts of raw observations can be structured as multidimensional data having time and space as their main dimensions. As a result, a semantic model for PA and PLF data should take into account similar challenges. Extensions of DCAT such as GeoDCAT and StatDCAT could contribute towards this direction.

The objective of this paper is to present the challenges for developing a semantic model for PA and PLF data discovery. The semantic model will be used to annotate metadata of both raw and processed PA and PLF data in order to facilitate their discovery and use them in advanced data driven application. For example enable queries such as Give me datasets of area X at the time frame [2018 - 2019] that contain data for soya yield.

The rest of this paper is structured as follows. Section 2 describes the approach that we follow. Section 3 briefly describes the most important data that can be found in PA and PLF scenarios, while section 4 presents the main roles that can be found in similar scenarios. Section 5 discusses challenges that are of statistical nature and section 6 introduces the semantic model. Finally, section 7 concludes our work and delineates future activities.

2   Methodology

The work towards the definition of the semantic model and the identification of the statistical challenges was conducted within the EU funded project Cybele[4] that aims to generate innovation and create value in the domains of agri-food by implementing Precision Agriculture and Precision Livestock Farming methods. Within the project we had the chance to deal with the needs of real pilots (e.g. soya yield prediction, sustainable pig production, aquaculture monitoring, open sea fishing) and the datasets they have. The methodology followed comprises the steps:

- Study the data used by the pilots. Details about the data are provided at section 3.
- Identify the user roles regarding data exploitation and their requirements. The roles and requirements are described at section 4.
- Extract the main concepts of the model from the requirements.
- Identify the statistical challenges that need to be addressed in order to define the semantic model (section 5)
- Define the semantic model by matching the concepts extracted from the requirements to existing standards and vocabularies (section 6)

3   Big agricultural and livestock farming data

Big agricultural and livestock farming data come from diverse sources and are available in different forms. Such data include:

- **Sensor data** are continuously collected through dedicated hardware (e.g. probes) and produce spatiotemporal measurements e.g. measure the soil's electrical conductivity at a specific location and time. Sensors produce large volume of data since measurements are repeated regularly (e.g. every 1 minute). However, usually aggregated data e.g. at level of day, is required in order to support decision making.
- **Earth observations** e.g. satellite images, drone aerial images, hyper-spectral images, RGB images. This type of data can produce huge volume of spatiotemporal data since they provide high resolution images of the earth. However, usually a single indicator (e.g. NDVI - Normalized Difference Vegetation Index) is required from each image in order to support decision making.
- **Video** e.g. video data from pig pens to monitor pigs behaviour. This type also produces huge volume of data. However, in this case only the identification of a specific behaviour (e.g. the pig drinks water) in time is actually required. This is also an aggregation of data and can be represented in a plain CSV file.

---

[4] www.cybele-project.eu

– **Crowd-sourced data** and human observations are collected through manual measurements and inspections (e.g. health inspection at livestock farms). Usually these are not of big volume, but need to be combined with other data e.g. sensor data, to support decision making.
– **Forecasts** e.g. for weather, prices, production. These data are also of spatiotemporal nature and usually are not of big volume. They can be combined with other data to facilitate decision making.
– **Maps** can be combined with other data to provide easily interpretable results and visualization e.g. show the NDVI on a map.

Although many types of agricultural and livestock farming data are of big volume (i.e. sensor and earth observations), what is usually needed for the decision making is aggreted data that provide an overview of data or useful indicators extracted from the data. Additionally, data is required to be combined e.g. visualize sensor data on maps. The combination requires the identification of compatible data e.g. data that have the same spatial coverage.

4 User Roles and Requirements

In order to identify the user requirements it is crucial to identify the user roles that will deal with the model and the expected uses of the model. The aim of the model is to facilitate data exploitation thus only roles regarding exploitation are encountered. These roles are:

– **End users**: exploit big data applications that produce easy to consume and interpret visualizations. This category includes for example farmers and livestock managers.
– **Modelers and developers**: produce big data driven application and models to be consumed by the end users.
– **Data analysts and farming consultants**: exploit data-driven decision making and support to farmers and livestock managers.
– **Statisticians**: exploit big agricultural and livestock farming data to deliver official statistics.

The user requirements are related to the exploitation of data and specifically to the data discovery and exploration. Table 1 present the semantic model requirements in terms of competency aspects related to data discovery, that should be considered at the design of the model i.e. what should the model be able to express. The requirements are expressed in the form "subject, predicate, object" (e.g. dataset is published by an organization) and are related to: i) provenance e.g. publisher, issuance date, ii) the theme of the dataset e.g. soya, iii) the spatial/temporal coverage of the dataset, iv) the activity that created the dataset, v) the structure of the dataset e.g. dimensions, measures, vi) the distribution e.g. format, license. For each of the requirements the table contains also the main model concepts that occur.

Table 1. Dataset discovery requirements and concepts

Requirement	Concepts
Search for datasets that belong at a repository	Dataset, Catalog
Dataset contains data about a specific cultivation (e.g. soya) or livestock	Dataset, Dataset, Theme (e.g. cultivation, livestock)
Dataset measures e.g. NDVI	Dataset, measurement
Dataset is published by an organization	Dataset, publisher
Dataset contains data that are in a specific language e.g. English	Dataset, Language
Dataset is issued/modified after/before e.g. 1/1/2019	Dataset, issuing/modification date
Dataset is updated e.g. monthly	Dataset, update frequency
Dataset contains data with temporal coverage e.g. [1/1/2017 - 31/12/2017]	Dataset, temporal coverage
Dataset contains data with temporal coverage e.g. [1/1/2017 - 31/12/2017]	Dataset, temporal coverage
Dataset contains measurements with temporal spacing e.g. one hour (measurements are repeated every one hour)	Dataset, temporal resolution
Dataset contains data with spatial coverage e.g. an area defined by a polygon	Dataset, spatial coverage
Dataset contains measurements minimum distance between items e.g. 30 meters	Dataset, spatial resolution
Dataset contains data about a specific theme e.g. weather data, price data	Dataset, theme
Dataset is the result of an activity that involves e.g. sensors, humans, satellites	Dataset, activity, agent (human, hardware)
Dataset is the result of an aggregation activity of other data (e.g. raw data)	Dataset, Activity
Dataset conforms to a model/schema e.g. SSN ontology	Dataset, standard
Dataset has specific dimensions e.g. time, geography	Dataset, dimension
Dataset uses a unit of measure e.g. prices in euro	Dataset, unit of measure
Dataset is distributed under a specific license	Dataset, distribution, license
Dataset is distributed in a specific format e.g. CSV, XML, Json	Dataset, distribution, format
Dataset is distributed through a data service e.g API	Dataset, distribution, data service
Dataset is accessed through a web page	Dataset, web page
Dataset distribution can be downloaded through a URL	Dataset, Distribution, Download URL
Data service is accessed through an endpoint URL	Data service, Endpoint URL

5 Semantic challenges

Raw PA and PLF data (e.g. sensor data) contain information at a fine grained level and a high volume that cannot be easily exploited. Usually only aggregated data are required in order to support decision making. As an example Santipantakis et al. [10] use semantic technologies to integrate big spatio-temporal data related to the mobility of entities (e.g. trajectories) by creating synopsis of the data and annotate them based on a common model. The synopsis are aggregations of data that contains annotation of critical points of the trajectory such as takeoff, landing etc.

At precision agriculture there is also a need for data that provide information at a higher level than the raw data. For example, at soya cultivation it is required to compute and incorporate indexes at parcel level such as the NDVI - Normalized Difference Vegetation Index, soil compression and water holding capacity. NDVI can be calculated using multispectral satellite images which contain red and infrared channels, while soil compression and water holding capacity can be produced by electrical conductivity probes (i.e. sensors) located at the parcels. The join between the two datasets can use as an ID the name of the parcel, the name of the farmer or the parcels location.

Additionally, in order to correlate the satellite and soil data with the agricultural practice and yields, we need to have the "ground truth". Ground truth can be collected either from crowdsourced data (e.g. information about yields, irrigation, fertilizers, pesticides, costs of operations) or from combine harvesters equipped with GPS trackers (e.g. yield maps, elevation maps). The join between satellite, soil data and the "ground truth" can use as ID the name of the parcel, the name of the farmer or the parcels location.

Similar cases occur also at precision livestock farming. For example, at pig farming it is crucial to measure and optimize pigs weight. For this reason, are required data about pig weights and pig pen conditions. Pig weight data can come as a result of processing raw video data or through human inspections. Additional data can come from sensors located at the pig pens e.g. pen humidity/temperature, water flow, fouling. Sensor data usually is needed to be aggregated e.g. at level of day while measurements occur every hour. The join between the two datasets can use as an ID the pig pen or the individual pig.

At the aggregation of data, the processing of data to calculate indexes (e.g. NDVI, pig weight) or the join between different data many challenges occur related to the definition of aggregation dimensions, measures, units of measure, aggregation functions and indexes. These aggregated data are in the form of data cubes, thus the challenges are also of statistical nature.

Another challenge that is important for both the raw and the aggregated data is the representation of the activity that generated the data. For example, raw data can be generated as a result of a sensor measuring activity, satellite imaging activity, human inspections etc. Aggregated data is the result of an aggregation activity on top of the raw data. The representation of this information is usefull for data provenance but also for the identification of data e.g. search for data

that are produced by sensors or search for data that measure NDVI as a result of satellite images processing.

6   Semantic model

This section presents existing vocabularies, ontologies, code lists (section 6.1) and matches them to the semantic model concepts (section 6.2).

6.1   Semantic Vocabularies and ontologies

This section presents a review of semantic vocabularies relevant to the precision agriculture and livestock farming. Two types of vocabularies were identified:

- Metadata vocabularies that enable the definition of metadata about datasets (e.g. geographical coverage)
- Domain vocabularies that can be used to populate the metadata values (e.g. specific geographical areas, livestock species). The domains of interest are the PA and PLF.

Meta-data Vocabularies   Data Catalog Vocabulary (DCAT)[5] is a W3C recommendation designed to facilitate interoperability between data catalogs published on the Web. By using DCAT to define metadata of data catalogs, publishers increase discoverability and enable applications easily to consume metadata from multiple catalogs. DCAT does not make any assumptions about the format (e.g. CSV, RDF, SQL) of the datasets described in the catalog. DCAT defines three main classes:

- *dcat:Catalog* represents the catalog that is a collections of metadata about datasets or data services.
- *dcat:Dataset* represents a dataset in a catalog. dataset in DCAT is defined as a "collection of data, published or curated by a single agent, and available for access or download in one or more formats".
- *dcat:Distribution* represents an accessible form of a dataset as for example a downloadable file.

A working draft of DCAT V2[6] introduces another class the *dcat:DataService* that represents a data service in a catalog. A data service is a collection of operations accessible through an interface (API) that provide access to one or more datasets or data processing functions.

DCAT defines diverse metadata properties including the data theme, the spatial/temporal coverage, the access rights, the license as well as information about the publisher, the publication date etc.

---

[5] https://www.w3.org/TR/vocab-dcat/
[6] https://www.w3.org/TR/2019/WD-vocab-dcat-2-20190528/

The ISA2 programme of European Commission has published DCAT-AP [4], a DCAT application profile for data portals in Europe. This application profile specifies the metadata records to meet the needs of data portals in Europe while providing semantic interoperability with other applications through the reuse of established controlled vocabularies (e.g. EuroVoc) and mappings to existing metadata vocabularies (e.g. Dublin Core, SDMX, INSPIRE metadata, etc.).

DCAT-AP has two extensions: i) GeoDCAT-AP[3] for describing geospatial datasets (extension includes concepts such as the spatial resolution, coordinate reference system) and ii) StatDCAT-AP[11] for statistical datasets (extension includes concepts such as the unit of measure, dimension/attribute properties).

RDF Data Cube (QB) Vocabulary[7] enables the publishing of multi-dimensional aggregated data, such as statistics, on the web. Data collected from sensors can easily be aggregated and expressed as multidimensional data[7]. This facilitates the applying of data analytics and visualizations on them. The QB vocabulary can be combined with DCAT in order to express statistical metadata of PA and PLF datasets.

The PROV-O ontology (PROV-O)[8] is a W3C recommendation, which describes provenance in terms of relationships between three main types of concepts:

– *prov:Entity*, which represents (physical, digital, or other types of) things;
– *prov:Activity*, which occur over time and can use and/or generate entities;
– *prov:Agent*, which are responsible for activities occurring, entities existing, or another agents activity.

PROV-O can be combined with DCAT in order to express provenance metadata of PA and PLF datasets.

Domain vocabularies Domain vocabularies can be used to populate the metadata values of precision agriculture and precision livestock farming datasets. The following paragraphs presents such domain vocabularies.

The Semantic Sensor Network (SSN) ontology[9] is a W3C recommendation for describing sensors and their observations. It supports many use cases, including satellite imagery, agriculture meteorology etc. SSN requires other ontologies to define domain semantics, units of measurement, time and location. Additionally, SSN uses the PROV-O ontology to represent the activity (e.g. observation) and the equipment (e.g. sensor) that created the data.

AGROVOC[10] is a controlled vocabulary covering areas related to the Food and Agriculture Organization of the United Nations including food, nutrition, agriculture, forestry, fisheries, scientific and common names of animals and plants, environment, biological notions, techniques of plant cultivation and more.

---

[7] https://www.w3.org/TR/vocab-data-cube/
[8] https://www.w3.org/TR/prov-o/
[9] https://www.w3.org/TR/vocab-ssn/
[10] http://aims.fao.org/vest-registry/vocabularies/agrovoc-multilingual-agricultural-thesaurus

A series of livestock ontologies[11] have been created by the French National Institute for Agricultural Research including the "Animal Trait Ontology for Livestock (ATOL)", the "Environment Ontology for Livestock (EOL)' and the "Animal Health Ontology for Livestock (AHOL)".

The "Quantity, Unit, Dimension and Type" (QUDT)[12] collection of ontologies facilitates the modelling of physical quantities and units of measure. A similar ontology is the Ontology of units of Measure (OM) 2.0 [9].

OWL-time ontology[13] is a W3C recommendation for describing the temporal properties of resources in any data. The ontology provides a vocabulary for expressing and sharing facts about topological (ordering) relations among instants and intervals. A similar vocabulary is proposed by reference.data.gov.uk[14]

INSPIRE[15] is an EU directive focusing at spatial data. It is based on ISO/OGC (ISO 19100 series) standard for geographical information. It addresses many themes including the "Agricultural and Aquaculture Facilities" [1]. INSPIRE has defined a code lists registry[16] related to farming (e.g. livestock species, soil types). FOODIE[8] is an ontology that extends INSPIRE data model for the publication of farm-related data (e.g. farm management) as Linked Data.

The W3C/OGC Spatial Data on the Web Interest Group has published "Best Practices on publishing Spatial Data on the Web"[17]. The same group is working on the "Statistical Data on the Web Best Practices". A preparatory work towards the best practice formulation has already been published by Kalampokis et. al [6].

6.2 Model definition and alignment

Based in the requirements collected (section 4) a model has been created (Figure 1). The central concept of the model is the *Dataset*. The *Dataset*:

 – is part of a *Catalog* that contains many datasets,
 – is published by a *Publisher* that can be a person or an organization
 – is available through a *Distribution* (e.g. download file, data service) and
 – is the result of an *Activity* that involves *Agents* (e.g. human sensor).

The *Dataset* has also properties including the *Theme* (a categorization of the data based on their domain), *Language*, *Issuing/Modification* date, *Update frequency*, *Measure*, *Unit of measure*, *Dimension*, *Spatial/Temporal coverage*, *Spatial/Temporal resolution*, *Standard* and *Web page*. Finally, the *Distribution* has properties including the *License*, *Format* and *Download URL*.

The semantic model concepts are mapped to concepts already defined at existing standard vocabularies. The vocabularies used for the matching are:

---

[11] http://www.atol-ontology.com
[12] http://www.qudt.org/
[13] https://www.w3.org/TR/owl-time/
[14] http://reference.data.gov.uk/def/intervals
[15] https://inspire.ec.europa.eu/
[16] http://inspire.ec.europa.eu/codelist
[17] https://www.w3.org/TR/sdw-bp/

- DCAT including some concepts from the ongoing work on DCAT v2. Prefix used `dcat:http://www.w3.org/ns/dcat#`
- StatDCAT extension of DCAT for statistical data. It is used for concepts of statistical nature such as the dimension and unit of measure (the property measure is not defined at the current version Stat DCAT 1.0.1). Prefix used `stat:http://data.europa.eu/statdcat-ap/`
- Dublic Core Metadata Initiative. It is used by DCAT for some concepts. Prefix used `dct:http://purl.org/dc/terms/`
- PROV Ontology[18]. Prefix used `http://www.w3.org/ns/prov#`
- The RDF Data Cube (QB) Vocabulary. It can be used to define the measure of the data. Prefix used `qb:http://purl.org/linked-data/cube#`
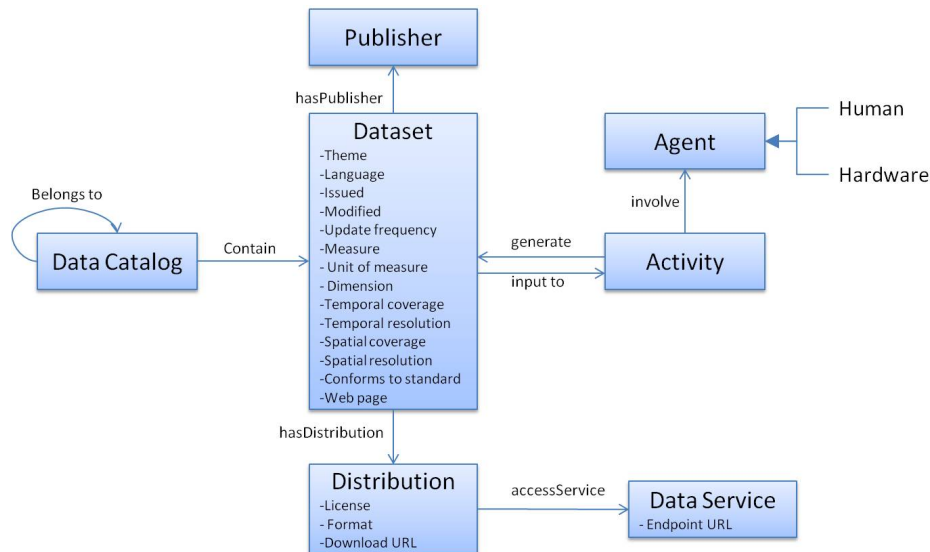

Fig. 1. Semantic Model for Precision Agriculture and Precision Livestock Farming

7 Conclusion

Precision Agriculture (PA) and Precision Livestock Farming (PLF) use massive amounts of data to improve crop production, animal health and welfare, and the environmental impact of agriculture and livestock production. The data that are used come from numerous and heterogeneous sources, such as satelites, drones, probes, local meteorological stations, and video recordings.

---

[18] https://www.w3.org/TR/prov-o/

Table 2. Mapping semantic model concept to vocabularies

Model concept	Mapping to vocabularies
Catalog	dcat:Catalog (Class)
Dataset	dcat:Dataset (Class)
Theme	dcat:theme (Property, domain → dcat:Dataset, range → skos:Concept)
Measurement	Not available. The qb:measure has as domain qb:ComponentSet and cannot be used at the dcat:Dataset.
Publisher	dct:publisher (Property, range → foaf:Agen
Language	dct:language (Property, range → dct:LinguisticSystem)
Issuing date	dct:issued (Property, range → xsd:date)
Modification date	dct:modified (Property, range → xsd:date)
Update frequency	dct:accrualPeriodicity (Property, range → dct:Frequency)
Temporal coverage	dct:temporal (Property, range → dct:PeriodOfTime)
Temporal resolution	dcat:temporalResolution (Property, range → xsd:duration)
Spatial coverage	dct:spatial (Property, range → dct:Location)
Spatial resolution	dcat:spatialResolutionInMeters (Property, range → xsd:decimal)
Activity	prov:Activity (Class), (dcat:Dataset, prov:wasGeneretadBy, prov:Activity)
Agent (human, hardware)	prov:Agent (Class), (prov:Activity, prov:wasAssociatedWith, prov:Agent)
Standard	dct:conformsTo (Property, range → dct:Standard)
Dimension	stat:dimension (Property, range → qb:DimensionProperty)
Unit of measure	stat:statUnitMeasure (Property, range → skos:Concept)
Distribution	dcat:Distribution (Class)
License	dct:license (Property, range → dct:LicenseDocument)
Format	dcat:mediaType (Property, domain → dcat:Distribution, range → dct:MediaTypeOrExtend)
Data service	dcat:DataService (Class)
Web page	dcat:landingPage (Property. range → foaf:Document)
Download UR	dcat:downloadURL (Property, domain → dcat:Distribution, range rdfs:Resource)
Endpoint URL	dcat:endpointURL (Property, domain → dcat:DataService, range: rdfs:Resource)

A common approach to address the challenge of discovering data across numerous heterogeneous sources is to semantically annotate and publish their metadata. This could enable users to search based on a standardised approach in a transparent way for them.

In the case of PA and PLF, however, the majority of the raw data is of fine granularity coming from sensors, satellites, or drones. This type of data usually needs to be further processed to produce metrics that will be used in data-driven decision making scenarios. A typical index that is widely used in PA is the

Normalized Difference Vegetation Index (NDVI), which quantifies vegetation. NVDI can be calculated from multispectral satellite images containing red and infrared channels. These new metrics could be of coarser granularity than the raw data depending on the problem at hand.

In some other cases, the raw data that are collected through various observation and sensing activities need to be aggregated before used in data-driven decision making scenarios. For example, raw meteo sensor data need to be aggregated in order to extract a valuable metric e.g. average temperature per day.

In this paper we described the challenges that are related to the multidimensional nature of PA and PLF data. We also proposed a semantic model that could address these challenges and facilitate the use of big data in PA and PLF scenarios. The semantic model re-uses existing vocabularies, however there are still some concepts that cannot be expressed. For example, there is no property to associate a dataset with the measures it contains. StatDCAT defines a property to associate only the dataset dimensions, while the QB vocabulary defines the property qb:measure that however is not applicable to dcat:Datasets. Additionally, there is no code list that can express all the data themes related to PA and PLF.

The final version of the proposed model will be used and evaluated in a real world settings to support PA and PLF applications in the CYBELE EU funded research project.

References

1. Inspire thematic wg agricultural and aquaculture facilities. d2.8.iii.9 data specification on agricultural and aquaculture facilities, december 2013
2. Berckmans, D.: Precision livestock farming technologies for welfare management in intensive livestock systems. Scientific and Technical Review of the Office International des Epizooties 33(1), 189–196 (2014)
3. Commission, E.: Geodcat-ap: A geospatial extension for the dcat application profile for data portals in europe version 1.0.1 (2016)
4. Dragan, A., Sofou, N.: Dcat application profile for data portals in europe version 1.2.1 (2019)
5. Harmon, T., Kvien, C., Mulla, D., Hoggenboom, G., Judy, J., Hook, J., et al.: Precision agriculture scenario. In: NSF workshop on sensors for environmental observatories. Baltimore, MD, USA: World Tech. Evaluation Center (2005)
6. Kalampokis, E., Zeginis, D., Tarabanis, K.: On modeling linked open statistical data. Journal of Web Semantics 55, 56 – 68 (2019), http://www.sciencedirect.com/science/article/pii/S1570826818300544
7. Lefort, L., Bobruk, J., Haller, A., Taylor, K., Woolf, A.: A linked sensor data cube for a 100 year homogenised daily temperature dataset. In: In Proceedings of the 5th International Conference on Semantic Sensor Networks. pp. 1–16 (2012)

8. Palma, R., Reznik, T., Esbr, M., Charvat, K., Mazurek, C.: An inspire-based vocabulary for the publication of agricultural linked data. In: Proceedings of OWLED 2015: Ontology Engineering, LNCS, vol 9557. pp. 124–133 (2016)
9. Rijgersberg, H., Van Assem, M., Top, J.: Ontology of units of measure and related concepts. Semantic Web Journal 4(1), 3–13 (2013)
10. Santipantakis, G., Glenis, A., Patroumpas, K., Vlachou, A., Doulkeridis, C., Vouros, G., Pelekis, N., Theodoridis, Y.: Spartan: Semantic integration of big spatio-temporal data from streaming and archival sources. Future Generation Computer Systems (In press) (2018)
11. Sofou, N., Dragan, A.: Statdcat-ap dcat application profile for description of statistical datasets version 1.0.1 (2019)