Analysis of Observational Variables from an Ontological Patterns Perspective

Natália Queiroz de Oliveira[1], Vânia Borges [1], and Maria Luiza Machado Campos [1]

[1]*Programa de Pós-Graduação em Informática (PPGI) – Universidade Federal do Rio de Janeiro (UFRJ) CEP-21941-901 - Rio de Janeiro – RJ - Brazil*

Abstract

An observational variable encodes what was measured, observed, derived, or computed in relation to Earth systems and phenomena representation in general. Well defined variables make data easier to find and reuse. However, increasing semantic interoperability of a variable associated concept is still a challenge. In order to avoid inconsistencies and ambiguities between different variable interpretations, it is essential to use a common terminology to homogeneously represent the core elements usually hidden in the variable description or naming. The Measurement Ontology Pattern Language (M-OPL) addresses the core conceptualization for measurements according to an ontology pattern language (OPL). It establishes standards for representing common core measurement concepts across various application domains. This paper discusses the use of M-OPL in the ontology of the I-ADOPT framework, promoting its semantic enrichment. As a result, we present an I-ADOPT alignment to the patterns established by M-OPL, with additional extension proposals to contemplate the particularities of the measurement domain.

Keywords

Observational Variables, Ontological Patterns, Measurement.

1. Introduction

Patterns are instruments for encapsulating common knowledge. The term “pattern language” in the Software Engineering community refers to a network of interrelated patterns together with a process for systematically solving coarse-grained software development problems [1]. This approach has been successfully exploited in Ontology Engineering with the development of ontology patterns (OPs). OPs are an emerging approach that benefits the reuse of encoded experiences and good practices [2], giving rise to ontology pattern languages (OPLs). Ontology Engineering is a complex task, considering the need for speedy development, motivating reuse in this area. However, an ontology engineer should also be careful with the complexity in precisely defining concepts and relations in an ontology.

An OP describes a particular recurring modeling problem that arises in specific ontology development contexts, presenting a well-proven solution for the problem [3]. OPLs provide guidance on how to reuse and integrate related patterns into a conceptual model and help an ontology engineer in selecting specific ontology patterns, depending on the problem being modeled according to a specific context. As a result, OPLs may produce gains in reuse and improve the quality of the resulting ontologies, as observed, for example, in OP initiatives developed by the research group Ontology and Conceptual Modeling Research Group (NEMO)[2]. This group has been working on many OPLs initiatives such as: Software Process OPL (SP-OPL) [4], ISO-based Software Process OPL (ISP-OPL) [5], Enterprise OPL (E-OPL) [6], Measurement OPL (M-OPL) [7], and Service OPL (S-OPL) [8].

Proceedings of the 15th Seminar on Ontology Research in Brazil (ONTOBRAS) and 6th Doctoral and Masters Consortium on Ontologies (WTDO), November 22-25, 2022
EMAIL: natalia.oliveira@ppgi.com.br (Natália Queiroz de Oliveira); vjborges30@ufrj.br (Vânia Borges); mluiza@ppgi.com.br (Maria Luiza Machado Campos)
ORCID: 0000-0001-8371-142X (Natália Queiroz de Oliveira); 0000-0002-6717-1168 (Vânia Borges); 0000-0002-7930-612X (Maria Luiza Machado Campos)
© 2022 Copyright for this paper by its authors.
Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).
CEUR Workshop Proceedings (CEUR-WS.org)

[2] https://nemo.inf.ufes.br/en/

CEUR-WS.org/Vol-3346/Paper5.pdf

The Measurement Ontology Pattern Language (M-OPL) addresses the main conceptualization associated with measurements in general, taking the Unified Foundational Ontology (UFO) as a basis [7]. Measurement is an essential tool of scientific investigation and discovery, and it enables complex phenomena of the universe to be precisely described. In technology, the increasing complexity and speed of many modern processes and machines make automatic control essential, and such control is not possible without satisfactory means of measurement [9] representation.

Measurement can be defined as a set of actions aiming to characterize an entity by attributing values to its properties [7]. Due to this definition, measurement can be applied in several domains as they share some concepts in common. Therefore, it is possible to identify core concepts that are independent of the application domain.

In order to avoid inconsistencies and ambiguities between different domains, it is important to use a common terminology to represent core concepts shared by them. Ontologies have been recognized as an important instrument for making knowledge clearer, promoting a common understanding, and avoiding inconsistencies and ambiguities between different domains. Through a shared conceptualization, ontologies can play the role of a "contract" established between parties for the purposes of communication and semantic interoperability [10].

A core ontology provides a precise definition of the structural knowledge in a specific domain that spans several application domains in that field [11]. Core ontologies are conceived aiming their reuse. By providing a network of patterns, an OPL improves the potential for reuse of a core ontology by enabling the selective use of parts of the core ontology in a modular and flexible way.

There are different meanings associated with the term "variable", depending on the context where it occurs [12]. From the Latin *variabilis*, a variable is that which varies or can vary. In research, variables are any measurable characteristics that can take on different values, qualities, traits or attributes of a particular individual, object, or situation being studied. Variables are commonly used in the biodiversity domain [13,14,15]. Those variables describing what was measured, observed, derived, or computed in relation to the Earth system are encoded as observational variables.

The InteroperAble Descriptions of Observable Property Terminology Working Group (I-ADOPT WG)[3] was responsible for creating a community-agreed framework for representing different aspects of observable variables like those in environmental research. The I-ADOPT framework ontology was designed to facilitate interoperability between existing variable description schemes (including domain-specific ontologies, semantic models and structured controlled vocabularies) [16]. A variable in the I-ADOPT is used as a synonym for an observable property as it is the description of something observed or derived.

In this work, we describe the use of M-OPL in the I-ADOPT framework ontology, first identifying concepts and relations that are semantically overloaded in the ontology. The goal is to semantically enrich the ontology, using M-OPL as a reference, to clarify core measurement concepts common in several application domains. By aligning the I-ADOPT ontology to the core modeling patterns proposed for measurements, we aim to capture the conceptualization of measurements for the compound concepts variables. Moreover, this alignment contributes to representing concepts not yet addressed by the ontology, such as scales and units, measurement procedures, measurement planning and measurement analysis.

This paper is organized as follows: section 2 presents the background of the paper, including a review of the I-ADOPT framework ontology and M-OPL. Section 3 describes the use of M-OPL, initially applying five pattern groups in the I-ADOPT ontology to clarify the conceptualization of measurements for compound concepts variables. Additionally, this section presents the relation between I-ADOPT ontology concepts, M-OPL patterns, and UFO fragments. Section 4 describes how M-OPL can be used to extend the I-ADOPT ontology concepts and relations. Finally, in section 5, we conclude and list some future work.

2. Background

[3] https://www.rd-alliance.org/groups/interoperable-descriptions-observable-property-terminology-wg-i-adopt-wg

In this section, we present an overview of the I-ADOPT framework ontology. The I-ADOPT WG was created in 2019 under the umbrella of the Research Data Alliance (RDA) Vocabulary Services Interest Group (VSSIG). In the meantime, we also present an overview of M-OPL pattern groups covering six measurement aspects besides a process suggesting an order to apply them.

2.1. I-ADOPT Framework ontology

The I-ADOPT WG had a strong focus on variables observed in environmental research as it leveraged existing efforts to accurately encode what was measured, observed, derived, or computed in relation to Earth systems [16]. This group was created to address the gap of deep metadata that further contextualize observations such as methodology, variables, and parameters. Those metadata currently vary from unstandardized free-text to controlled vocabularies such as Climate and Forecast Standard Names[4] or the British Oceanographic Data Centre (BODC) Parameter Usage Vocabulary[5].

The development of the I-ADOPT framework followed a bottom-up approach through phases. The initial phase was dedicated to the collection of user stories from the environmental domain, identifying key requirements, and analysing existing semantic presentations of scientific variables and terminologies in use. The proposed framework was then tested against a variety of examples to ensure that it could be used as a sound basis for the creation of new variable names as needed. Finally, the results were formalised into the I-ADOPT ontology and subsequently extended in an output of recommendations guideline [17].

The I-ADOPT ontology was inspired by the atomization approach of the "Complex Property Model" [18] and the "Scientific Variables Ontology" [19]. This approach conceives the Variable as a compound concept consisting of at least one entity (ObjectOfInterest) and one Property. In addition, other entities can be included to help contextualize the target object of observation. Although the Scientific Variables Ontology is intrinsically terminology agnostic, in some cases, especially for human accessibility, it may be necessary to identify a concept with a standard label, providing a snapshot of the information associated with a particular compound concept variable [19].

Figure 1 shows the schema and instance levels overview of the proposed I-ADOPT framework ontology [20]. The schema level comprises four main classes (Variable, Property, Constraint, and Entity) and six relations (hasProperty, hasObjectOfInterest, hasContextObject, hasMatrix, hasConstraint, constrains). The ontology has been defined as variable-centric as the variable is a complex semantic representation of any type of data acquisition event, be it a human-based observation, a sensor-based measurement, a calculation, or a simulation [16].

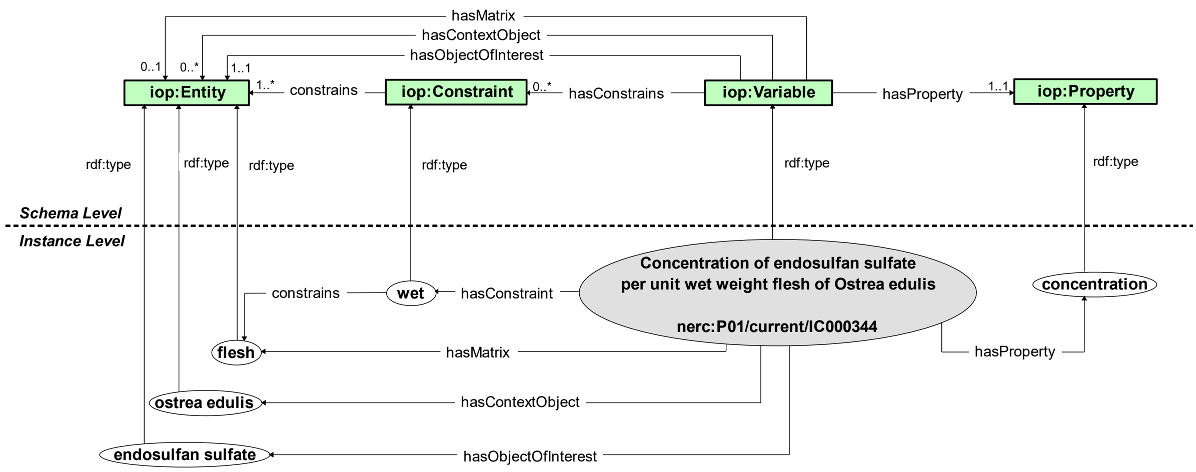

Figure 1: The I-ADOPT framework ontology - schema and instance levels overview

To explain the concepts and relations at the instance level, we use an example of a complex biodiversity compound concept variable, adapted from [16]. We even considered aspects that were

[4] https://cfconventions.org/standard-names.html
[5] https://www.bodc.ac.uk/resources/vocabularies/parameter_codes/

missing in the original I-ADOPT concept discussion, according to the NERC Vocabulary Server (NVS). The concept is "*Concentration of endosulfan sulfate per unit wet weight flesh of Ostrea edulis*" which refers to the quantitative result (i.e., requiring a magnitude and unit) of a measurement. It is important to mention that in [16] per unit and weight aspects of the compound concept variable were not considered, but for the schema and instance levels overview proposed in Figure 1, we must include them to be faithful to the concept.

As aforementioned, a Variable is a description of something observed or derived as a compound concept, consisting of at least one entity (the ObjectOfInterest) and its Property. The Property (concentration) is a type of characteristic of the ObjectOfInterest. The Entity is an object or occurrence that has a role in an observation. An Entity may play one of the following roles: ObjectOfInterest (endosulfan sulfate), ContextObject (ostrea edulis), or Matrix (flesh). The Constraint (wet) limits the scope of the observation and restricts the context to a particular state. It describes relevant properties of the involved entities in the particular observation. These concepts are interconnected using the following object properties:

- **hasProperty**: It relates a Property with a Variable, with a cardinality of 1..1. This cardinality indicates that the Variable has exactly one Property;
- **hasObjectOfInterest**: It associates the Variable with the ObjectOfInterest, i.e., the Entity whose property is observed. Similar to the previous one, its cardinality is 1..1, meaning that a Variable requires exactly one ObjectOfInterest;
- **hasContextObject**: It associates the Variable with entities that provide additional information regarding the ObjectOfInterest, i.e., ContextObject entities. Its cardinality is 0..*, which means that a Variable may have more than one Entity associated in this context or none;
- **hasMatrix**: It associates the Variable with the Matrix in which the ObjectOfInterest is contained. It is not mandatory, and when it exists, it should only show one Matrix, so its cardinality is 0..1;
- **hasContraint**: It relates to the constraints associated with a Variable, being optional. Its cardinality is 0..*;
- **constrains**: It associates a Constraint with an Entity of the Variable. A Constraint can constrain one or more Entities. Its cardinality is 1..*.

It is important to highlight that in the I-ADOPT ontology, entities assume a role by means of the relation they are associated with. Consequently, the same entity can appear as ObjectOfInterest, ContextObject, or Matrix. Therefore, it could also have different roles depending on the particular variable. Besides, the ontology does not yet cover any additional concept or relation associated with units, instruments, methods, time-related and geographical location information. Units are essential information for describing measures, but a quantitative variable might be expressed differently, requiring units to be modeled independently of variables. Although these concepts provide essential information for interpreting actual observations, they were not originally intended to be included in the scope of the I-ADOPT framework. For this reason, in Figure 1, the context information of "per unit weight" could not be associated with any concept or relation at the schema level of the ontology.

2.2. Measurement Ontology Pattern Language (M-OPL)

The M-OPL was developed following a pattern-oriented design approach. It addresses the core conceptualization for measurements, according to UFO, and consists of six modules covering the following aspects [7]: (i) Measurement Entities, including MEnt and TMElement patterns, related to entities and their properties that can be measured; (ii) Measures, dealing with Mea and TMea patterns, defining measures and classifying them according to their dependence on other measures; (iii) Measurement Units & Scales, contemplating MScale, TMScale, MUnit, MUnit & Scales patterns, concerning the scales related to measures and the measurement units used to partition scales; (iv) Measurement Procedures, considering MProc, TMProc, MProcBM, MForm, MProcDM patterns, dealing with the procedures needed to collect data for measures; (v) Measurement Planning, including TMGoal, INeed, MPI, MPI-MP and Ind patterns, addressing the goals that drive measurement as well as the measures used to verify goals achievement; and finally (vi) Measurement & Analysis, dealing with Meas and MAna patterns, concerning data collection and analysis.

Following the OPL approach [7], these modules compose a catalog of patterns to be adopted and associated with a process, suggesting their application order, as presented in Figure 2. The M-OPL patterns are presented below according to each relevant aspect of measurement they address.

MEnt and TMElement patterns are defined to address aspects concerning the definition of measurable entities and properties. The former handles entities and their measurable elements identification. The second pattern defines the measurable element type, i.e., whether the element is directly or indirectly measurable. Directly measurable elements do not depend on others to be measured, such as the body weight. Indirectly measurable elements, on the other hand, depend on other measurable elements, such as the velocity of a body, which depends on the distance traveled and time.

MScale, TMScale, MUnit and MUnit&Scale patterns deal with aspects of scales and units of measurement associated with measures. MScale defines the scale and the constituent values. TMScale pattern establishes the scale types, while MUnit pattern defines the associated Measurement Unit. Finally, MUnit&Scale represents the relationship between units and scales.

MProc, TMProc, MProcBM, MForm, MProcDM patterns address aspects concerning measurement procedures. These procedures describe the steps to collect data for the measures. TMProc is applicable in ontologies that address different types of measurement procedures. MProcBM and MProcDM are patterns for defining procedures for base measures and derived measures respectively. MForm is employed with MProcDM, dealing with measurement formulas to calculate derived measurements.

TMGoal, INeed, MPI, MPI-MP and Ind patterns treat aspects related to measurement planning. TMGoal defines measurement goal and its subdivisions, comprising simple or composed measurement goals. INeed pattern deals with the necessary information identified from the measurement goals. MPI pattern handles measurement planning, specifying which measurement should be collected for each goal or required information. MPI-MP pattern is responsible for choosing the measurement procedure to be used. Ind pattern establishes the measurement that works as an indicator to evaluate the achievement of some measurement goals.

Meas and MAna patterns are concerned with modeling aspects related to Measurement and Analysis. The Meas pattern address the data collection, defining the Measurement concept and its Relations. The MAna pattern is responsible for the analysis.

Figure 2 presents the process proposed in [7], represented using an UML activity diagram, suggesting an order to apply M-OPL. Patterns are presented in gray and the darker lines the recommended paths. This process is used in the discussion of the next sections (3 and 4).

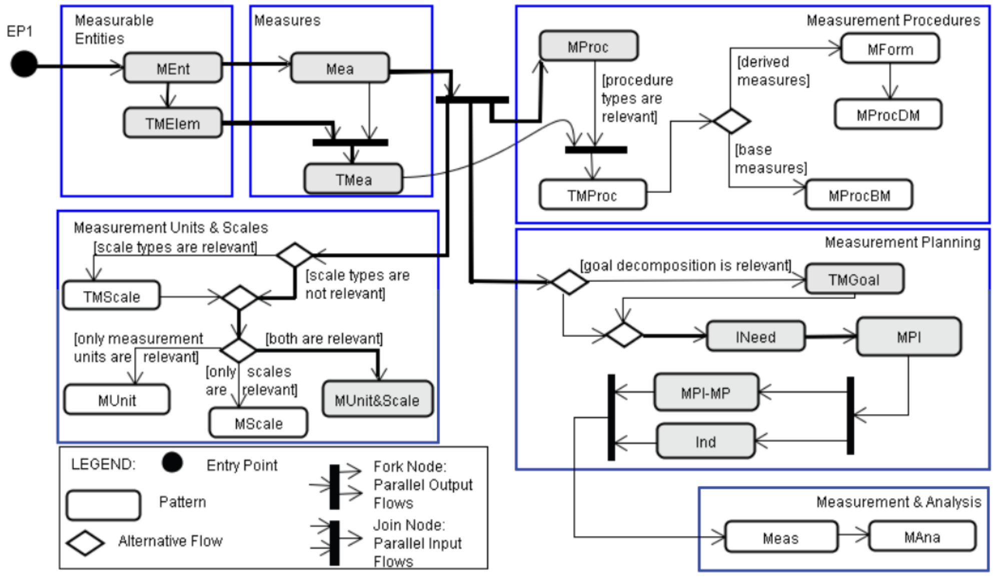

Figure 2: M-OPL suggested process [7]

In general, an OPL specification includes, for each pattern, some competency questions (CQs) which the pattern aims to answer. Therefore, CQs can be reused when a pattern in OPL is selected. For example, once a Domain-Related Ontology Pattern (DROP) is chosen, its CQs can be extended for the domain ontology [3]. This reuse helps to improve the productivity of the Ontology Engineering process [21]. According to [22] a set of CQs, based on Measurable Entities & Measures sub-ontology, was used to evaluate a Software Measurement Ontology (SMO). Some of them were reused in this work in order to align the I-ADOPT ontology to the core modeling patterns proposed for measurements, as the following:

Q1. Which is the measurable entity? *
Q2. What is the type of a measurable entity?
Q3. Which are the measurable elements of a measurable entity?
Q4. Can the measurable element be directly measured? **
If yes, Q4.1 Which measurable elements can be directly measured?
Q5. From which other measurable elements can an indirectly measurable element be measured?
Q6. Which are the measurable elements that characterize a measurable entity?
Q7. Which measures can be used to quantify a measurable element?
Q8. Which measures should be measured to compute a derived measure?
Q9. What is the measure unit of the measure?
Q10. What is the scale of the measure?
Q11. What is the type of a scale?
Q12. Which are the values of a scale?
Q13. Which are the measurement procedures that are applied to a measure?
Q14. What are the calculation formulas used in a measurement procedure?
Q15. Which information have to be met by measures in order to monitor the measurement goal? *
Q16. Which are the measurement analysis procedures that are applied to a measure?
Q17. Which measures are correlated to a given measure?

** CQ created for the related domain*
*** CQ adapted from [22]*

In the literature, we found related works using M-OPL in different domains. In [7] the first version of M-OPL was proposed and it was used to build a Software Measurement Ontology (SMO), which aims to capture the conceptualization involved in this domain, including traditional and high maturity aspects of software measurement. In [23], the authors describe the use of M-OPL in the scenario of measurements associated with performance monitoring of Internet links, generating a new version of an ontology developed in the context of the Pinger-LOD Project [24]. Another work [25] presents an ontology that provides a way towards capturing and leveraging the intensity of Beliefs, Desires, Intentions and Feelings (BDIFs), during a Knowledge-intensive Process (KiP) execution. It was built based on the M-OPL, the Speech Act Theory, and the Knowledge-intensive Process Ontology as a formal conceptualization to measure BDIFs in KiPs.

Here, we present a new domain for using M-OPL, the domain of observational variables. Because these variables, observed in the biodiversity field, require a complex semantic representation to describe the data acquisition and prevent ambiguous descriptions. Like many other domains, biodiversity research has been transformed by a big data revolution, where providing information interoperation is urgently required to support responses towards a sustainable future.

3. Use of M-OPL in I-ADOPT ontology

In this section, we present how M-OPL can contribute to clarify and capture the conceptualization of measurements for compound concepts variables in the I-ADOPT ontology, and also more specific aspects concerning measurement in high maturity representation levels. M-OPL ontology patterns were developed taking UFO as a basis. Hence, we also present how the I-ADOPT ontology concepts and relations can be aligned to M-OPL patterns and UFO fragments.

The competency questions presented in section 2.2 were used, as they play a prominent role in defining the scope and purpose of the domain conceptualization, serving as a testbed for ontology evaluation. We also used the same concept example, obtained from NVS, to exemplify the instances addressed to M-OPL patterns.

3.1. Application of MEnt, TElem, Mea, TMea, and MUnit pattern groups

In order to use M-OPL, we initially selected five pattern groups (MEnt, TElem, Mea, TMea, and MUnit) to be applied, based on the sequence suggested by the process presented in Figure 2. Figure 3 shows a fragment of the application of those M-OPL pattern groups as a proposal to capture the conceptualization of measurements for compound concepts variables. The M-OPL patterns are depicted in gray, followed by the instances of the concept in white. The concept example is illustrated in orange, and it lines emphasize the I-ADOPT ontology relations.

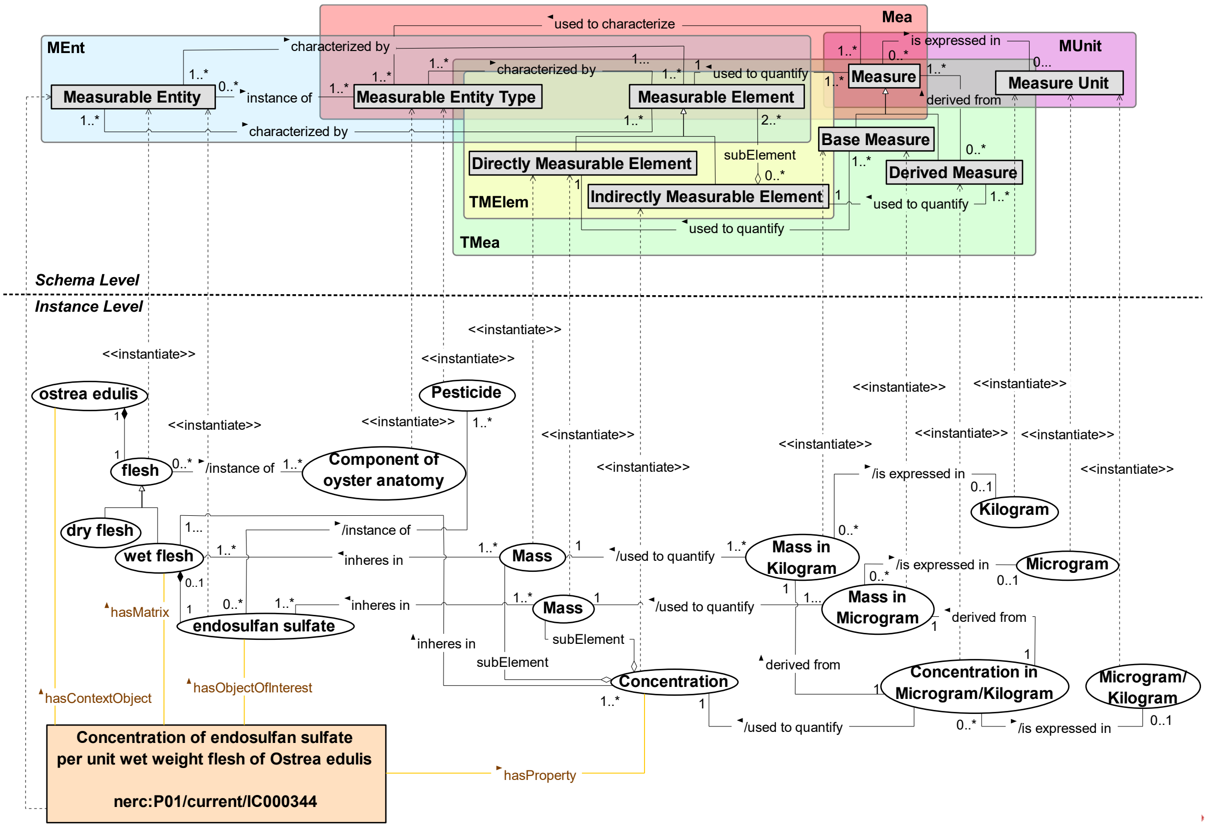

Figure 3: Application of MEnt, TElem, Mea,TMea, and MUnit pattern groups

The first pattern applied was MEnt (Measurable Entity), that has been instantiated in order to consider the Measurable Entity for the compound concept variable, answering competency question Q1: *wet flesh*. Measurable entities are classified according to Measurable Entity Types, that also have been instantiated in order to consider the type of measurable entity relevant to the *Component of oyster anatomy* domain, answering competency question Q2.

After applying the pattern MEnt, two patterns were applied in parallel: TMElem (Types of Measurable Elements) and Mea (Measures). Through the pattern TMElem, we could identify the Measurable Element for the compound concept variable, answering competency question Q3: *Concentration/Mass (enfosulfan sulfate)/Mass (wet flesh).* It has been characterized as an Indirectly Measurable Element, as *Concentration* depends on sub-elements in order to be measured, addressing competency question Q5. According to NVS, concentration concept definition is "*the amount of a specified substance in a unit amount of another substance or matrix*". The sub-elements to measure Concentration have been instantiated as Directly Measurable Element (elements that do not depend on others to be measured), answering competency question Q4.1: *Mass (enfosulfan sulfate)/Mass (wet flesh).*

It is important to highlight that the relation "characterized by" between Measurable Entity and Measurable Element is specialized from the homonym relation between Universal and Moment Universal, according to UFO. Thereby the directly measurable element (*Mass*) and the indirectly measurable element (*Concentration*) characterize the Measurable Entity (*wet flesh*) as the directly measurable element (*Mass*) characterizes the measurable entity (*endosulfan sulfate*), answering competency question Q6.

In other words, after applying the pattern TMElem, we could also identify different measurable levels for the compound concept variable used as an example, due to the relation between the patterns Measurable Entity and Measurable Element. First, we have identified *wet flesh* as the only measurable entity and then we instantiated it also to consider *endosulfan sulfate* as another measurable entity, representing a whole-part relationship between its parts. With this in mind, the competency question Q1 answer was updated to: *wet flesh/endosulfan sulfate*, as measurable entities for the compound concept variable. Likewise, Q2 answer was also updated to*: Component of oyster anatomy/Pesticide*.

In the Mea pattern, Measures are used for quantifying Measurable Elements and to characterize Measurable Entity Types. Measure is a Function in the sense that it maps an instance of Measurable Element to a value. TMea pattern characterizes a Measure into two types: Base Measure, which is functionally independent of other measures and used to quantify Directly Measurable Elements, and Derived Measure, which is defined as a function of other measures and used to quantify Indirectly Measurable Elements. Hence, *Mass in Kilogram* and *Mass in Microgram* were instantiated as Base Measure, answering competency question Q7, as they are used to quantify the directly measurable elements (*flesh Mass* and *endosulfan sulfate Mass*). Likewise, *Concentration in Microgram/Kilogram* is an instance of Derived Measure that quantifies the indirectly measurable element (*Concentration*), answering competency question Q5. Besides, the Derived Measure *Concentration in Microgram/Kilogram* is derived from the Base Measures *Mass in Kilogram* and *Mass in Microgram*, answering Q8.

In the Measurement Units & Scales group, important for the domain in order to model the variable units, the pattern MUnit was applied. As aforementioned, a quantitative variable might be expressed in different ways which requires units to be modeled independently of the associated variables. As units are essential information for describing measures, these concepts contribute for the interpretation of actual observations. In NVS, the concept "*Concentration of endosulfan sulfate per unit wet weight flesh of Ostrea edulis*" has as related concept "*Micrograms per kilogram*" that indicates the unit of weight of flesh wet mass and endosulfan sulfate mass, addressing Q9.

According to the application of these patterns' groups, we emphasize how I-ADOPT ontology concepts and relations are semantically overloaded due to variable entities assuming different roles by the means of the relations they were associated with, e.g., ObjectOfInterest, ContextObject, or Matrix, and also due to the overload of the variable measurable entity and measurable element. Through the use of M-OPL, guided by the pattern's competency questions, it is possible to clarify the core measurement conceptualization hidden for a compound concept variable, as identifying measurement entities and their properties that can be measured, besides defining measures and classifying them according to their dependence on others measures. Beyond being used to identify these core measurement concepts, they can be reused in several domains as they share some concepts in common, reinforcing that M-OPL patterns can contribute to this.

3.2. Relation of I-ADOPT concepts, M-OPL and UFO

According to I-ADOPT, an observable variable, as a concept that provides metadata for values made available in datasets, is a compound of at least one entity representing the ObjectOfInterest and one Property. These elements are represented in Figure 4 as Observable Quantitative Variable, ObjectOfInterest Entity and Property, respectively. Figure 4 highlights the relation between I-ADOPT ontology concepts, M-OPL patterns and UFO fragments.

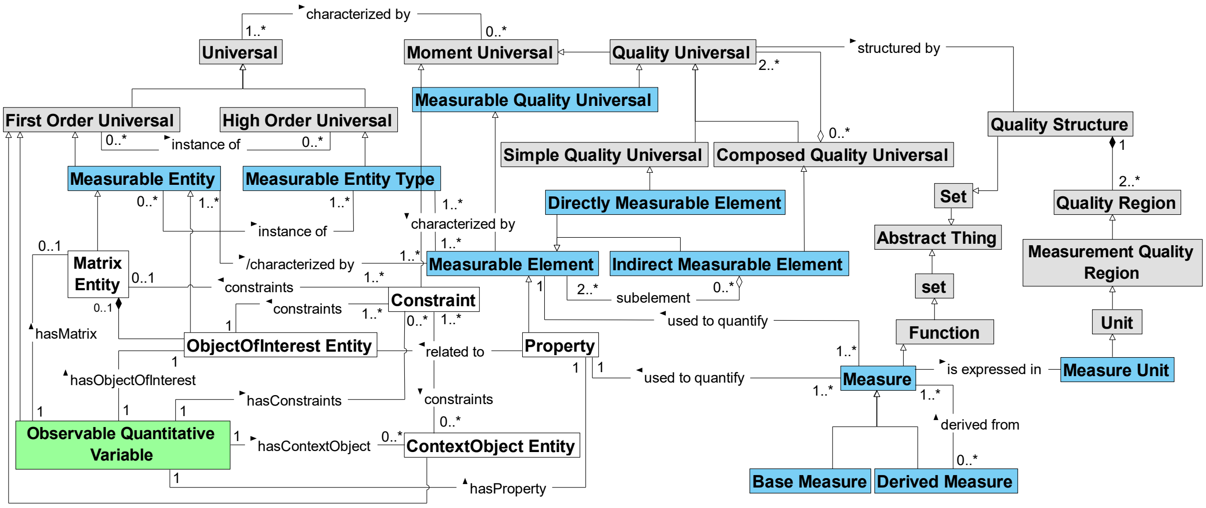

Figure 4: Relation of I-ADOPT concepts, M-OPL and UFO

The Property in a quantitative variable corresponds to the observed aspect, so it specializes the Measurable Element. The ObjectOfInterest Entity being the observable element to which the Property refers, specializes the Measurable Entity. The relate to relation is defined to represent the association of the Property with the ObjectOfInterest Entity. In addition to these elements, the Variable can optionally contain a Matrix, one or more ContextObject and one or more Constraint. These elements are represented in Figure 4 as Matrix Entity, ContextObject Entity and Constraint, respectively.

The Matrix represents the entity in which the ObjectOfInterest is contained. This information allows establishing a part-whole relation between Matrix Entity and ObjectOfInterest Entity whenever the variable presents a Matrix. Considering the part-whole relation, it also specializes Measurable Entity since it presents a measurable part. ContextObject refers to entities that provide context to the observation. Thus, in our model, ContextObject Entity is treated as a specialization of First Order-Universal.

The relations established by the I-ADOPT ontology are also represented in Figure 4. Thus, the constrains relation associates Constraint with Matrix Entity, ObjectOfInterest Entity and ContextObject Entity. These restrictive aspects (Constraints) can define a specialization on the entity they constrain: the specialization of Matrix – *flesh* in *wet flesh* specifies exactly the flesh type considered by the observation.

The relations hasMatrix, hasObjectOfInterest, hasContextObject, hasConstraint and hasProperty are also considered. These relationships associate the Observable Quantitative Variable with its constituent elements. Intuitively, Observable Quantitative Variable emerges from its association with the other concepts in the ontology. In UFO, it could be considered a *Moment Universal Relator*. It represents the mereological sums of two or more externally dependent modes, i.e., of aspects of other individuals [23]. It represents a reification mechanism of the quantitative variable observation event. For our model, however, this concept specializes a First Order Universal.

The relation between the measurable entity and the measurable element (Property) needs the context of the observation. In the instantiation shown in Section 3.1, it is observed that the Property (*concentration*) is a measurable element of the Matrix (*wet flesh*), derived from the measurable elements of the Matrix Entity and the ObjectOfInterest Entity.

By aligning I-ADOPT concepts with M-OPL and UFO fragments, it is possible to: (i) treat the semantic overload in the Entity concept, making explicit the interactions between them, e.g., the Matrix Entity contains the Object of Interest Entity; (ii) add new concepts referring to measurement, such as Indirect and Directly Measurable Element; (iii) provide new properties associating the established concepts to Functions, such as Measure used to quantify Property, and to Quality Regions, defining the Measure Unit in which the Variable Measure should be expressed. It is worth mentioning that these new elements are part of an ontological pattern based on UFO ontology, specifically acting with measurement aspects, regardless of the domain.

4. Use of M-OPL to extend the I-ADOPT ontology

In this section, we present how M-OPL can be used to extend the I-ADOPT ontology concepts and relations, also capturing the conceptualization of measurement units & scales, measurement procedures, measurement planning, and measurement & analysis. In doing so, the I-ADOPT ontology can take advantage of these conceptualizations, semantically enriching the ontology schema level, to represent other important core measurement aspects from the compound concept variable, addressing complementary competency questions.

Figure 5 shows a fragment of the application of M-OPL pattern groups to extend the I-ADOPT ontology. It contemplates, besides the application of five M-OPL pattern groups presented in gray, other pattern groups highlighted in orange.

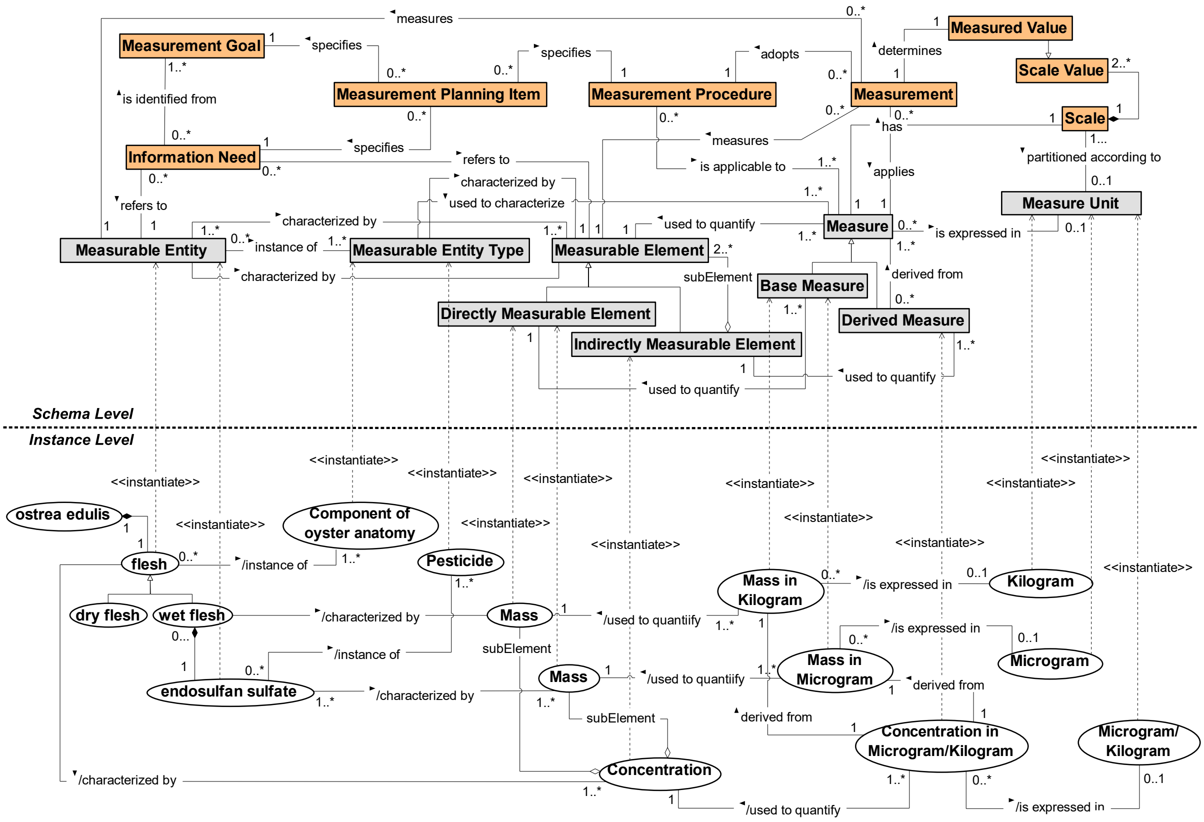

Figure 5: Extension of I-ADOPT ontology through M-OPL use

Considering the Measurement Units & Scales pattern group, the pattern MUnit&Scale is applied as it is also important to consider the variable scales of measures in the domain. According to MUnit&Scale pattern, measures can be expressed in Measure Units in which a measure is expressed as partitions of its scale. Measures have Scales composed of all possible values (Scale Value) to be associated by the measure to a measurable element. Reusing this pattern, the CQs 10, 11 and 12 could be addressed.

In the Measurement Procedures group, a Measured Procedure is applicable to a Measure. The MProc pattern is applied to model the steps to be carried out aiming a data collection for measures, addressing the CQs13 and 14. In the Measurement Planning group, the first applied pattern was INeed as it corresponds to information needs identified from measurement goals, addressing the CQ 15.

Measurement Goals are targets that can be used to guide the identification of the measures needed in a certain context. The next applied pattern was MPI-MP (Measurement Planning Item – Measurement Procedure) as it connects a Measurement Goal, an Information Need, a Measure, and a Measurement Procedure, meaning that the Measure meets the Information Need that is identified from the Measurement Goal.

Finally, the Measurement & Analysis group was applied. In the Meas (Measurement) pattern, it is possible to model data collection and analysis. Measurement is performed based on a Measurement Planning Item and it measures a Measurable Element of a Measurable Entity by applying a Measure and adopting a Measurement Procedure. The result is a Measured Value, which refers to a value of a measure scale. These could address CQs 16 and 17.

In summary, M-OPL contributes to address the conceptualization associated with observational variable measurement as it is organized in extensible modules according to specific measurement contexts. It can guide a common way of decomposing compound concept variables, identifying and reusing essential components as measurement entities, measures, measurement units & scales, measurement procedures, measurement planning and measurement & analysis.

5. Conclusion and Future Works

OPLs facilitate the reuse of integrated ontological patterns into a conceptual model, leading to gains in reuse and the resulting ontologies quality. M-OPL addresses the core conceptualization for measurements in general and their characterization, organized according to an OPL. Measurement is very common in several domains as they share some concepts in common. An observational variable is a complex semantic representation of any type of data acquisition that usually carries ambiguous descriptions in human readable form or even for a machine. Especially for machine accessibility, it may be necessary to identify a concept with a standard label, providing a snapshot of the information associated with a particular compound concept variable.

The I-ADOPT framework ontology conceives a variable as a compound concept consisting of at least one entity (ObjectOfInterest) and one Property. Beyond that, other entities can be included to help contextualize the target object of observation. Although the ontology does not yet cover any additional concept or relation associated with units, instruments, measurable methods, time-related and geographical location information, these concepts provide essential information for interpreting actual observations. In particular, Units are essential information for describing measures and a quantitative variable might be expressed differently, requiring units to be modeled independently of variables.

This paper presented the use of M-OPL as a standard proposal for capturing the conceptualization of measurements for compound concepts variables in the I-ADOPT ontology. It also highlighted more specific aspects concerning measurement in high maturity representation levels. M-OPL addresses main measurements conceptualization that can be applied in several domains, as they share some concepts in common, contributing to create and to reuse complex unambiguous variable descriptions in a machine and human readable form. The benefits of using M-OPL to the I-ADOPT ontology are: (i) identification of concepts and relations semantically overloaded, since M-OPL patterns have been developed following a largely explored theory based on UFO; (ii) alignment to the core modeling patterns proposed for measurement, common to several application domains; (iii) capture of the conceptualization of measurements for compound concepts variables; and (iv) an extension to represent concepts not yet addressed by the ontology as scales and units, measurement procedures, measurement planning and measurement analysis. Besides the highlighted benefits, the I-ADOPT ontology could take advantage of the semantic richness of M-OPL to guide a common way of decomposing compound concept variables, identifying and reusing essential components as measurement entities, measures, measurement units & scales, measurement procedures, measurement planning and measurement & analysis.

As future work, we intend to use ontological patterns to represent qualitative variables and also explore new uses of M-OPL aiming to increase semantic interoperability of a larger set of complex concept variables. We believe this will contribute to avoiding inconsistencies and ambiguities between different variable interpretations, identifying main core concepts that are independent of the application domain.

6. Acknowledgements

This work has been partially supported with student grants from CAPES (Process numbers 223038.014313/2020-19 and 88887.613048/2021-00).

7. References

[1] F. Buschmann, K. Henney, and D. Schimdt. Pattern-oriented Software Architecture: On Patterns and Pattern Language, volume 5. John Wiley & sons, 2007.

[2] R. Falbo, G. Guizzardi, A. Gangemi, and V. Presutti, Ontology patterns: clarifying concepts and terminology, in: Proceedings of the 4th Workshop on Ontology and Semantic Web Patterns (Vol. 1188).

[3] R. Falbo, et al, Ontology pattern languages, in: Ontology Engineering with Ontology Design Patterns, IOS press, 2016, pp. 133-159.

[4] R. Falbo, M. Barcellos, J. Nardi, and G. Guizzardi, Organizing ontology design patterns as ontology pattern languages, in: Extended Semantic Web Conference, Springer, Berlin, pp. 61-75.

[5] F. Ruy, R. Falbo, M. Barcellos, G. Guizzardi, and G. Quirino, An ISO-based software process ontology pattern language and its application for harmonizing standards, ACM SIGAPP Applied Computing Review, 15(2):27–40, 2015.

[6] R. Falbo, F. Ruy, G. Guizzardi, M. Barcellos, and J. P. Almeida, Towards an enterprise ontology pattern language, in: Proceedings of the 29th Annual ACM Symposium on Applied Computing, pp 323–330, ACM, 2014.

[7] M. Barcellos, R. Falbo, and V. Frauches, Towards a measurement ontology pattern language. in: ONTO.COM/ODISE FOIS, 2014.

[8] R. Falbo, et al, An ontology pattern language for service modeling, in: Proceedings of the 31st Annual ACM Symposium on Applied Computing. ACM, 2016.

[9] L. Finkelstein. Fundamental concepts of measurement. ACTA IMEKO, v. 3, n. 1, pp. 10-15, 2014.

[10] G. Guizzardi, Ontology, ontologies and the “I” of FAIR. Data Intelligence, v. 2, n. 1-2, pp. 181-191, 2020.

[11] A. Scherp, et al, Designing core ontologies, in: Applied Ontology, vol. 6, no. 3, pp. 177-221, 2011.

[12] A. Schoenfeld, and A. Arcavi, On the meaning of variable..The mathematics teacher, v. 81, n. 6, pp. 420-427, 1988.

[13] W.D Kissling, et al, Towards global data products of essential biodiversity variables on species traits, Nat. Ecol. Evol. 2, 1531–1540, 2018. doi: 10.1038/s41559-018-0667-3

[14] C. König, et al, Biodiversity data integration—the significance of data resolution and domain, PLoS Biol 17(3): e3000183, 2019. doi:10.1371/journal.pbio.3000183

[15] A. Hardisty, et al, The Bari Manifesto: An interoperability framework for essential biodiversity variables. Ecological informatics, v. 49, p. 22-31, 2019. doi:10.1016/j.ecoinf.2018.11.003

[16] B. Magagna, et al, The I-ADOPT Interoperability Framework for FAIRer data descriptions of biodiversity, 2021.

[17] B. Magagna, et al, InteroperAble Descriptions of Observable Property Terminologies (I-ADOPT) WG Outputs and Recommendations (1.1.0), 2022. doi:/10.15497/RDA00071.

[18] A. Leadbetter, and P. Vodden, Semantic linking of complex properties, monitoring processes and facilities in web-based representations of the environment, International Journal of Digital Earth, 9(3), pp. 300-324, 2016. doi:10.1080/17538947.2015.1033483.

[19] M. Stoica, and S. Peckham, An Ontology Blueprint for Constructing Qualitative and Quantitative Scientific Variables, in: ISWC (P&D/Industry/BlueSky), 2018.

[20] I-ADOPT Framework ontology. URL: https://i-adopt.github.io/index.html

[21] F. Ruy, C. Reginato, V. Santos, R. Falbo, and G. Guizzardi, Ontology engineering by combining ontology patterns, in: Proceeding of the 34th International Conference on Conceptual Modeling (ER’15), p. 173–186. Springer, 2015.

[22] M. Barcellos, R. Falbo, and R. Dal Moro, A well-founded software measurement ontology, in: Formal Ontology in Information Systems, pp. 213-226, IOS Press, 2010.

[23] R. P. Nunes, et al, Measurement ontology pattern language applied to network performance measurement, in: Brazilian Seminar on Ontological Research (Ontobras 2015), São Paulo, Brazil, 2015.

[24] R. F. Souza, et al, Linked Open Data Publication Strategies: Application in Networking Performance Measurement Data, in: The Second ASE International Conference on Big Data

Science and Computing, Stanford, ASE BIGDATA / SOCIALCOM / CYBERSECURITY Conference, 2014.
[25] P. H. P. Richetti, and F. A. Baião, A Measurement Ontology for Beliefs, Desires, Intentions and Feelings within Knowledge-intensive Processes, ONTOBRAS, 2018.
[26] C. Fonseca, et al, Relations in ontology-driven conceptual modeling, in: International Conference on Conceptual Modeling, Springer, Cham, p. 28-42, 2019.