Measurement Ontology Pattern Language Applied to Network Performance Measurement

Raphaela P. Nunes[1], Adriana S. Vivacqua[1], Maria Luiza M. Campos[1], Ana Carolina Almeida[2]

[1]Programa de Pós-Graduação em Informática - Universidade Federal do Rio de Janeiro (UFRJ)
Caixa Postal 68.530 – 21941-590 – Rio de Janeiro – RJ - Brasil

[2]Instituto de Matemática e Estatística - Departamento de Informática e Ciências da Computação - Universidade do Estado do Rio de Janeiro (UERJ)
Rua São Francisco Xavier, 524 - 6º andar - 20550-013 - Rio de Janeiro - RJ - Brazil

{mluiza,raphaela.nunes}@ppgi.ufrj.br, avivacqua@dcc.ufrj.br, ana.almeida@ime.uerj.br

Abstract. *The Measurement Ontology Pattern Language (M-OPL) addresses the measurement core conceptualization according to an Ontology Pattern Language (OPL). An OPL provides holistic support for solving ontology development problems for a particular field and guiding the development of ontologies. This paper presents the application of M-OPL in a specific domain, network performance measurement. As a result of this application, a discussion of the use of M-OPL is presented together with some suggestions of extensions to contemplate the peculiarities of this domain.*

1. Introduction

Measurement is a very important discipline in several domains, since it provides useful information for getting conclusions and making decisions [BARCELLOS *et al.*, 2014]. Measurement is the process of assigning numbers or symbols to properties of real-world entities, according to widely-defined rules, in order to describe them [FINKELSTEIN; LEANING, 1984]. It can also be understood as a process that involves a set of actions in order to characterize entities assigning values to their properties [BARCELLOS *et al.*, 2014].

When analyzing different areas where measurements can be applied, it is possible to identify some particular concepts related to the knowledge treated on each specific area. However, it is also possible to identify some core concepts that are independent of the application domain. In order to homogeneously represent these core concepts across different domains, avoiding inconsistencies and ambiguities, it is important to use a common terminology shared by the domains. Currently, core ontologies have been used to promote this common conceptualization.

The Measurement Ontology Pattern Language (M-OPL) [BARCELLOS *et al.*, 2014] addresses the main conceptualization associated to measurements in general, organized according to an Ontology Pattern Language (OPL). An OPL [FALBO *et al.*, 2013] corresponds to a network of interconnected ontology modeling patterns that provides holistic support for solving ontology development problems for a particular field.

In this work, we describe the application of M-OPL to the scenario of measurements associated to performance monitoring of Internet links. The objective is to discuss the scope and usage of M-OPL to generate a new version of the original ontology developed in the context of the Pinger-LOD Project [SOUZA *et al.*, 2014]. By aligning it to the core modeling patterns proposed for measurements, we aim to decrease the possibility of inconsistencies and ambiguities on the ontology and facilitate future publication and linkage of data with other related data sources in the Web.

During the application of M-OPL we are also concerned to represent multidimensional aspects of measures, to support the representation of different perspectives of a measurement. As in other domains, a proposed extension to an existing OPL may derive, in the future, new modeling patterns to be incorporated in a new version of the OPL, in this case, an extended version of M-OPL.

The remainder of this paper is organized as follows: Section 2 presents a brief description of M-OPL. In Section 3, we describe the application process of M-OPL to the network measurement domain, making firstly a brief explanation of the current structure of PingER ontology and how it was derived. In Section 4, related works and further discussions are presented. Finally, in Section 5, we conclude and list some future work.

2. Measurement Ontology Pattern Language (M-OPL)

A core ontology provides a precise definition of the structural knowledge in a specific field that spans across several application domains in that field [SCHERP *et al.*, 2012]. Ontology Pattern Languages (OPL) have been proposed and used to organize core ontologies facilitating their reuse and extension [FALBO *et al.*, 2013]. An OPL [FALBO *et al.*, 2013] provides a set of interconnected ontology modeling patterns and a process that describes how to combine them to build an ontology applied to a specific domain.

The M-OPL addresses the core conceptualization for measurements and their characterization. M-OPL includes six patterns, defined according to the Unified Foundational Ontology (UFO) [GUIZZARDI, 2005] and covering six measurements aspects: Measurement Entities, which include patterns related to the entities and their properties that can be measured; Measures, which deal with the definition of measures and classify them according to their dependence on other measures; Measurement Units & Scales, which concerns the scales related to measures and the measurements units used to partition the scales; Measurement Procedures, which deals with procedures required to collect data for measures; Measurement Planning, which addresses the goals that drive measurement and the measures used to verify goals achievement; and Measurement & Analysis, which concerns data collection and analysis.

In the application of M-OPL discussed in this article, we will be particularly interested in applying the Measures aspects, in order to define the measures associated to the network links measurement domain and characterize the main entity to be measured (Measurable Entity) in the field, which, in this case, is the Internet Link. For the purpose of this paper, we are not including details on measurement procedures or data collection.

3. M-OPL Application to Network Performance Measurement

In this section, we discuss how M-OPL was used to derive a new version of an ontology for the conceptualization of measurements for network links performance. The original ontology was developed in the context of the PingER (Ping End-to-end Reporting) project, which is conducted by the Network and Telecommunications Department at the SLAC[1] National Accelerator Laboratory, in Stanford University, USA. The project manages data about the quality of Internet links from 1998 to the present day, on an hourly and daily basis, comprising 16 different metrics collected by 80 monitor nodes over 800 monitored nodes (more than 8000 pairs of nodes), in more than 160 countries [COTTRELL, 2001]. Each measurement is basically defined by a ping sent from a monitor node to a monitored node at any given time, and related to a specific network metric, considering data packets sizes of 100 and 1000 bytes.

3.1. Original PingER Ontology

The original PingER ontology [SOUZA *et al.*, 2014] was developed to serve as a reference vocabulary and structure to represent and annotate PingER data as RDF triples for a linked data publishing and querying application.

The PingER ontology is an adaptation of the MOMENT (Monitoring and Measurement in the Next Generation Technologies) ontology [SALVADOR. *et al.*, 2010; RAO, 2010], a core ontology which conceptualizes the networking performance measurements domain.

The MOMENT ontology is complex and generic in the way it contemplates the main characteristics referring to network measurement. This generality of the ontology enables it to be adapted to many different network measurement scenarios, including the PingER domain. However, since the ontology is so generic, the ontology fails in representing PingER reality. Additionally, the ontology does not aim to minimize the number of triples generated, which make it harder to process a large amount of data. Thus, it was decided not to reuse the ontology exactly as it is, but, instead, to reuse its concepts and ideas as basis to build an ontology more specialized for the PingER domain, which could better support data analytical processing. The current version of Pinger ontology has been implemented and used to publish PingER data, that can be accessed from a SPARQL endpoint[2].

Figure 1 shows an overview of the generated model. In the center of the ontology is the main superclass, which is the *Measurement* class, representing the process of acquiring measures. *Measurement* relates to the following classes, in order to qualify the measurement: *Metric*, through *measuresMetric* relation to specify which network metric is being measured; *MeasurementParameters* which can be specialized in *PacketSize*, through *hasMeasurementParameters* relation, to specify the measurement attributes; *DateTime* through *hasDateTime* relation, to specify the time interval in which the

[1] https://www6.slac.stanford.edu/

[2] http://pingerlod.slac.stanford.edu/sparql

measurement was made; *SourceDestinationNodes,* which represents the Internet Links and is related to two types of *Network Nodes*, the one which performs the role of monitor node, sending the ping signal, and the other which performs the role of monitored node, receiving the ping. The relation is made through *PingER-ont:hasSourceNode* and *PingER-ont:hasDestinationNode* relations, respectively.

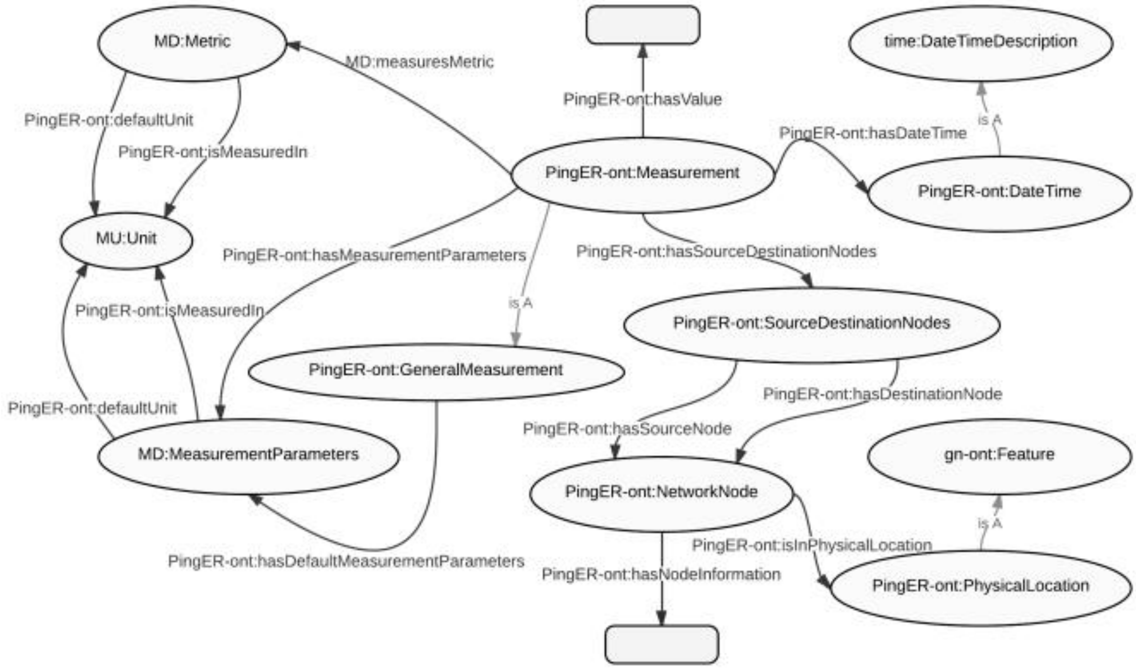

Figure 1. PingER Ontology (SOUZA *et al*., 2014)

To define the parameters of time (when the measure was taken) and space (where the network nodes - NetworkNode - are located), concepts extracted from Time [HOBBS & PAN, 2006] and Geonames [VATANT & WICK, 2012] ontologies were used.

3.2. M-OPL Application to PingER Network Performance Measurement

In order to apply M-OPL to the network performance measurement domain, we used the patterns depicted in gray in Figure 2, applied in the order indicated by the darker lines. The process was defined in the paper that originally presented M-OPL [BARCELLOS *et al*., 2014]. Figure 3 shows a fragment of the resulting ontology.

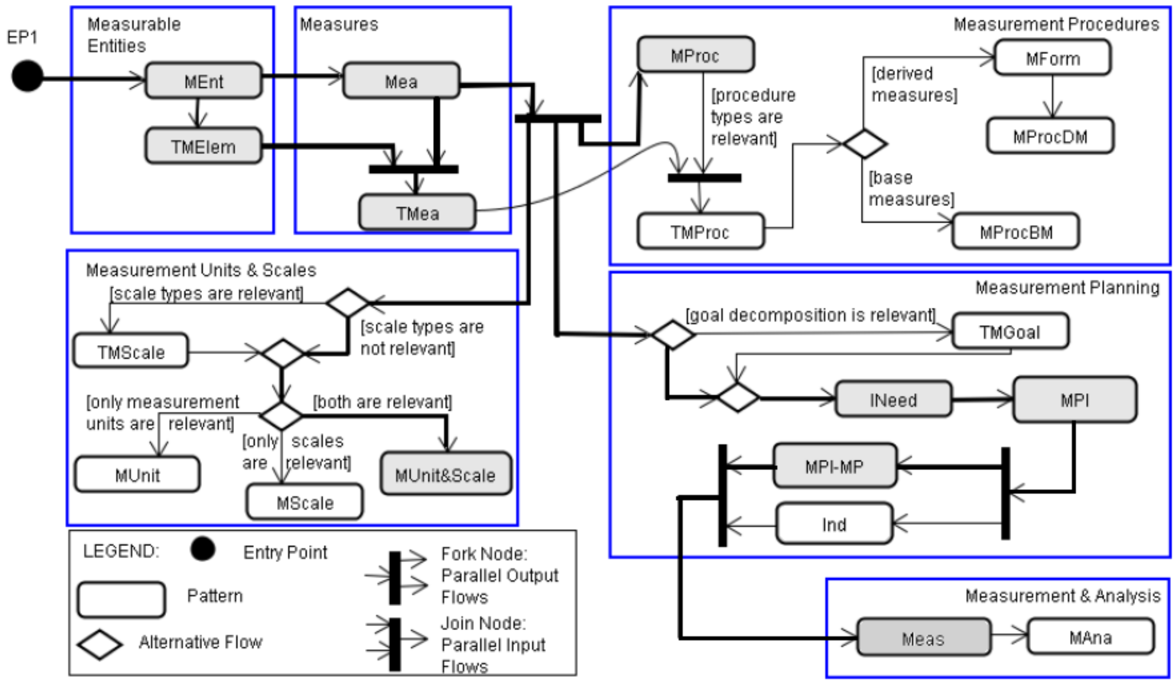

Figure 2. Application order of modeling patterns

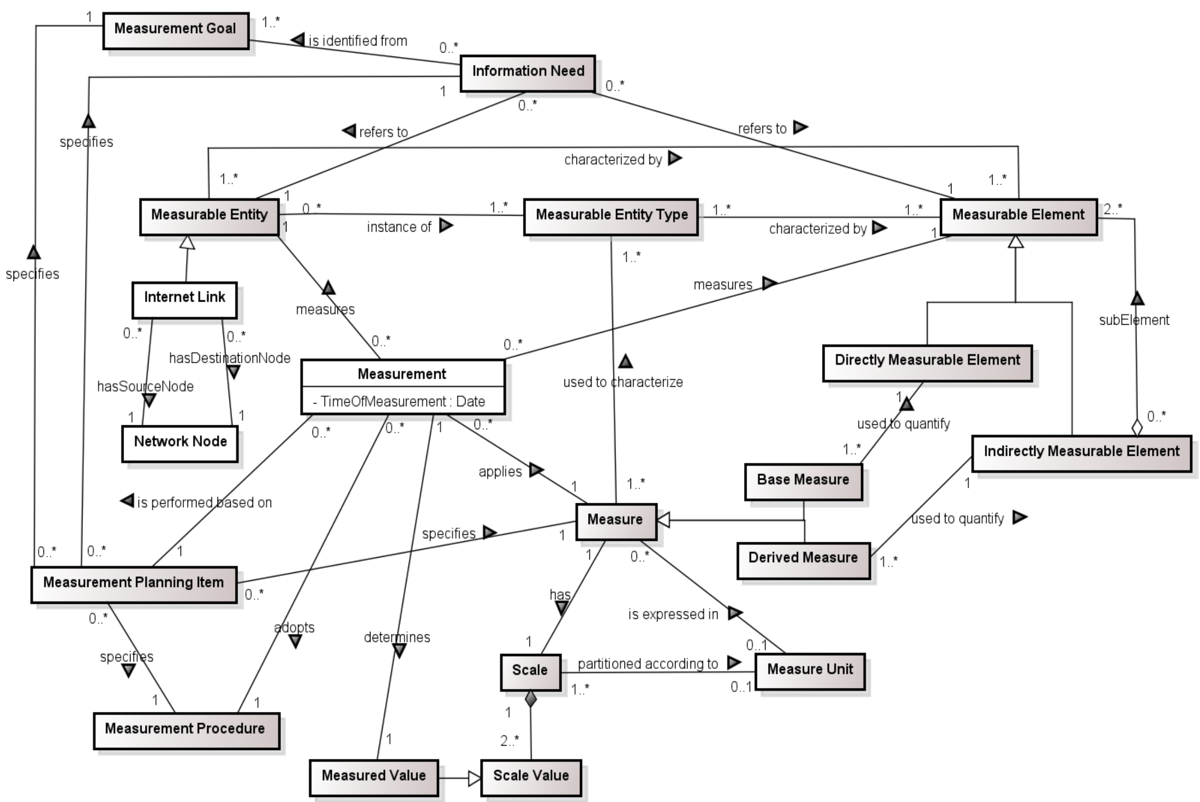

Figure 3. Derived Ontology

The first pattern applied was **MEnt** (Measurable Entity), which has been extended to consider the type of measurable entity relevant to the domain, an I*nternet link*. This is the current Measurable Entity Type being monitored by the PingER project, but others could be considered and then included.

After using the pattern **MEnt**, two patterns were applied: **TMElem** (Types of Measurable Elements) and **Mea** (Measures). In pattern **TMElem**, we could identify the Measurable Elements considered by the PingER ontology structure and characterize them as Directly Measurable Elements (elements that do not depend on others to be measured) or Indirectly Measurable Elements (elements that depend on other sub-elements to be measured).

Examples of Directly Measurable Elements are Duplicate Packets and Packet Loss, as they result from counting the associated events. Indirectly Measurable Elements depend on sub-elements to be measured, and in PingER case they include Directivity, Minimum Round Trip Delay, Conditional Loss Probability, Mean Opinion Score and Average Round Trip Time, among others. Round Trip Time or RTT, for example, is related to the distance between the nodes plus the delay at each hop along the path between them.

The **Mea** pattern was used as it was defined in M-OPL. In this pattern, a Measure quantifies a Measurable Element, characterizing a Measurable Entity Type. Hence, the measure *number of packets* quantifies the measurable element *packet loss* that characterizes the measurable entity of type *Internet Link*. But to better define an Internet Link it was necessary to extend the M-OPL, adding the concept of *NetworkNode*, which was related to *Internet Link* through *hasSourceNode* and *hasDestinationNode* relations, employed according to the role that the node is performing during measurement.

As the pattern **TMElem** was used, the pattern **TMea** (Type of Measures) also had to be used to characterize a Measure as a Base or Derived Measure, which serves to quantify Directly and Indirectly Measurable Elements, respectively. This pattern was also applied exactly as it was defined in M-OPL.

After using the **Mea** pattern, three paths were followed in parallel. The first led to the Measurement Units & Scales group, the second to the Measurement Procedures group and the third to the Measurement Planning group. In the Measurement Units & Scales group, as it is important for the domain in order to model the units and scales of measures, the pattern **MUnit&Scale** was used, as it was defined in M-OPL.

In the Measurement Procedures Group, as it is not important for the domain to detail the data collection procedures according to the different types of measures, only was used, in this group, the pattern **MProc**, as it was defined in M-OPL.

In the Measurement Planning Group, the first pattern used was **INeed**. For example, in our case, *Know the variability of service* could be considered as an instance of Information Need. It could be used to indicate the achievement of the Measurement Goal, which was defined in PingER domain as *Check the network quality*. Although not represented in the fragment of Figure 4, measurement goals may be composed or simple. In this case, *Check the network transfer capacity* could be a sub-goal of the composed measurement goal *Check the network quality*.

The next pattern used after **INeed** was **MPI-MP** (Measurement Planning Item – Measurement Procedure), which was applied in the same way that it was defined in the M-OPL. Finally, after addressing a Measurement Planning Item, the **Meas** (Measurement) pattern of Measurement & Analysis group was applied. In the **Meas** pattern, Measurement is executed based on a Measurement Planning Item and adopting a certain Measurement Procedure. It measures a Measurable Element of a Measurable Entity applying a Measure. The result is a Measured Value, which refers to a value of a measure scale.

To be able to represent temporal aspects in **Meas** pattern, *TimeOfMeasurement* was added as property of the relator Measurement (actually a property of the Event giving rise to the Relator). In UFO, Relators are derived from Events, which are temporal based constructs.

Making a brief comparison between the ontology derived from the application of M-OPL and the original PingER ontology, it is possible to note that the original version of PingER ontology is more focused on treating the particular domain concepts, representing only partially the semantics of measurements. M-OPL includes a general knowledge about measurements, applicable to different situations. By applying the M-OPL, these generic classes can be specialized according to the situation being considered. For example, in our scenario, the main focus of network measurements was performance evaluation, considered as quality measures associated to the network (using the Ping procedure as in the SLAC laboratory). However, by specializing M-OPL classes, we can use further grouping of goals and measures, and represent network evaluations other than related to performance/quality, like network reliability measures (which would include Medium Time Between Failures – MTBF, Gracefull Degradation, Recovery Time after Failures, Medium Time to Repair – MTTR, among others) [BALTRUNAS et al, 2014].

Considering the main patterns proposed in M-OPL and comparing with the classes in original structure of PingER ontology, it is possible to note that some of these patterns are already somehow represented in the original ontology structure. The Metric class is similar to the Types of Measurable Elements pattern (**TMElem**), since it is possible to represent the Measurable Elements through this class. But is not possible to distinguish the Measurable Elements which depend or not on others to be measured through this class, then it was necessary to add these new concepts in the ontology, by adding the Directly Measurable and Indirectly Measurable Elements.

The Unit class is similar to the Measurement Units and Scales pattern (**MUnit&Scale**), since it is possible to represent the units in which measures are expressed through this class. But it is not possible to represent in the original ontology the scales for measures which are partitioned according to the units, so it was necessary to create a new class to represent the Scale element and relate it with the Measure Unit class.

The Measurement class is similar to the Measurement pattern (**Meas**), since it functions, like in M-OPL, as a Relator, connecting the classes involved in the measurement process. However, the measurement process of the original ontology does not consider a Measurement Planning Item neither a Measurement Procedure, so it was necessary to create new classes to represent these elements and relate them with the Measurement class.

By using a generic conceptualization, the derived ontology allows interoperation with other complementary domains, which is particularly interesting for Semantic Web applications and publication of LOD.

4. Related Work

In the literature, two works were found applying the M-OPL. In the original proposal of M-OPL [BARCELLOS *et al*., 2014], it was used to build a Software Measurement Ontology (SMO), with a very straightforward application of the patterns. In [FRAUCHES, 2014], knowledge about the measuring process and the vocabulary adopted in the process described in M-OPL were used, as a basis for defining an approach for obtaining indicators from open data. The approach proposes a set of activities that must be performed from established measurement goals, to organize the data from an open database and to extract indicators that provide useful information for decision making.

Applications of OPL in different domains have also been presented, most of them confirming the possibility of reuse of the proposed patterns, and the usefulness of the accompanying guiding process for their application. There is, though, still a lack of cases where the derived domain ontologies have been implemented and where the patterns have been directly imported and adapted using an existing modeling tool. The reuse of model fragments is already supported by the OLED[3] (OntoUML Light Editor), but currently focusing on general patterns and anti-patterns included in its underlying library.

During the application of **Meas** in the PingER domain it was not evident how to explicitly represent the dimensions that qualify a measure, such as the time dimension and geographic location dimension, which could facilitate the visualization of possible analytical perspectives. But, in fact, considering M-OPL and its domain ontology derivations, it does not seem reasonable to contemplate a multidimensional structure, similar to the representation of the Data Cube Vocabulary [CYGANIAK; REYNOLDS; TENNISON, 2014], where facts (measurements) and dimensions (associated concepts) are at the core of the model. Although recognizing the long-term importance of multidimensional models for analytical processing (and, of course, for exploration and aggregation of measurements or statistical data) they serve a different purpose: to make explicit the analytical possibilities associated to the data. But, as such, this type of representation do not constitute a real conceptual model associated to a domain, as it does not usually support the representation of existing relations among concepts, their interdependences and other rules that constitute the rich semantics of the real world conceptualizations.

From the solution found to represent the time and location aspects associated with Measurement in PingER domain, it is possible to conclude that M-OPL can

[3] https://code.google.com/p/ontouml-lightweight-editor/

represent the temporal aspect, treating it as property of the Measurement Relator. But for other characterizations related to Measurement, it seems a better solution to represent them as new concepts, extending some of the patterns proposed in M-OPL.

5. Conclusion

In this paper, we have presented the application of M-OPL to the network performance measurement scenario, in order to derive a new version of PingER ontology [SOUZA *et al.*, 2014] so that we could take advantage of the semantic richness of measurements and their associated concepts when interoperating PingER data with other data as linked open data on the Web.

The benefits of applying M-OPL brought to the development of the new version of PingER ontology were: (i) decreasing the possibility of inconsistencies and ambiguities, since the basic patterns of the M-OPL have been developed following a largely explored theory based on UFO; (ii) acceleration of the ontology development process, as the patterns application process has proven to be effective and easy to use; (iii) as already stated previously, alignment to the core modeling patterns proposed for the measurement area can facilitate the future publication and linkage of the PingER data with related data sources in the Web.

As future work, we are already experimenting with this new version of the ontology, and we expect to evidentiate that the addition of semantic expressiveness brought to the model can lead to more sophisticated and intelligent applications, compensating the inherent increase of complexity of the ontology structure. Also, further discussion on the multidimensional characteristics of measures would be interesting, and what would be the best representation or derivation from a rich conceptualization such as the one already contemplated by M-OPL constructs.

References

Baltrunas, D., Elmokashfi, A., Kvalbein, A. (2014) "Measuring the Reliability of Mobile Broadband Networks". In: Proceedings of the 2014 Conference on Internet Measurement Conference, p. 45-58, New York, NY, USA.

Barcellos, M. P., Falbo, R. A., Frauches, V. G. V. (2014) "Towards a Measurement Ontology Pattern Language". In: 1st Joint Workshop ONTO.COM/ODISE on Ontologies in Conceptual Modeling and Informations Systems Engineering co-located with 8th International Conference on Formal Ontology in Information Systems, Rio de Janeiro, Brazil.

Cottrell, L. (2001) "Internet End-to-end Performance Monitoring", Available in: http://www-iepm.slac.stanford.edu. Last Access: July 2015.

Cyganiak, R., Reynolds, D., Tennison, J. (2014) "The RDF Data Cube Vocabulary". In: W3C Recommendation. Available in: http://www.w3.org/TR/vocab-data-cube/. Last Access: July 2015.

Falbo, R. A., Barcellos, M.P., Nardi, J. C., Guizzardi, G. (2013) "Organizing Ontology Design Patterns as Ontology Pattern Languages". In: 10th European Semantic Web Conference - ESWC 2013 - Semantics and Big Data, Montpellier. The Semantic Web: Semantics and Big Data - LNCS, 2013, v. 7882. p. 61-75, Springer.

Finkelstein, L., Leaning, M. S. (1984) "A Review of the Fundamental Concepts of Measurement". In: Measurement, v. 2, issue 1, p. 25-34, January-March 1984.

Frauches, V. G. V. (2014) "Uma Abordagem Baseada em Ontologias para Obtenção de Indicadores a partir de Dados Abertos", MSc. Dissertation, Department of Computer Science, Federal University of Espírito Santo, Vitória, ES, Brazil.

Guizzardi, G. (2005) "Ontological Foundations for Structural Conceptual Models", PhD Thesis, Centre for Telematics and Information Technology (CTIT), no 05-74, The Netherlands.

Hobbs, J. R., Pan, F. (2006) "Time Ontology in OWL". In: W3C Working Draft, v. 27, p. 133, September 2006.

Rao, S. (2010) "Monitoring and Measurement in the Next Generation Technologies". Available in: ftp://ftp.cordis.europa.eu/pub/fp7/ict/docs/future-networks/projects-moment-factsheet_en.pdf..

Salvador, A., de Vergara, J. L., Tropea, G., Blefari-Melazzi, N., Ferreiro, Á., Katsu, Á. (2010) "A Semantically Distributed Approach to Map IP Traffic Measurements to a Standardized Ontology", In: International Journal of Computer Networks & Communications (IJCNC), v. 2, no. 1, p. 13-31.

Scherp, A., Franz, T., Saathoff, C., Staab, S. (2012) "A Core Ontology on Events for Representing Occurrences in the Real World", In: Multimedia Tools and Applications, v. 58, issue 2, p. 293-331.

Souza, R. F., Cottrell, L., White, B., Campos, M. L., Mattoso, M. (2014) "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.

Vatant, B., Wick, M. (2012) "Geonames ontology", Available in: http://www.geonames.org/ontology/documentation.html, July 2015.