

# Redefinition and Statistical Analysis of Measures for Evaluating the Quality of Ontologies

Melina Tibaldo<sup>1</sup>, Alexia Wilkinson<sup>1</sup>, Ma. Laura Taverna<sup>1</sup>, Mariela Rico<sup>1</sup>, and Ma. Rosa Galli<sup>2</sup>

<sup>1</sup> Centro de Investigación y Desarrollo de Ingeniería en Sistemas de Información (CIDISI) - Universidad Tecnológica Nacional - Facultad Regional Santa Fe, Lavaise 610 - S3004EWB - Santa Fe - SF - Argentina  
[mrico@frsf.utn.edu.ar](mailto:mrico@frsf.utn.edu.ar)

<sup>2</sup> INGAR-UTN-CONICET, Avellaneda 3657, S3002GJC Santa Fe, Argentina

**Abstract.** OntoQualitas is a framework to evaluate an ontology whose purpose is the interchange of information between different contexts. However, the framework does not propose acceptance thresholds of the measure values. In this paper, measures proposed in this framework are redefined in order to improve their usefulness in assessing the quality of such ontologies. These measures were calculated semi-automatically on a set of ontologies and its results were described by means of a statistical analysis as a first step to the definition of their acceptance thresholds.

**Keywords:** ontology quality, measure, statistical analysis

## 1 Introduction

Even after more than a decade since the emergence of ontologies in Computer Science and with its growing use in different disciplines, standardized methods have not been developed for evaluating their quality [8].

Although methodologies, methods, techniques, and software tools to support the ontology building process were proposed, ontology evaluation still plays only a passive role in ontology engineering projects [17]. In order to assess the ontology quality, different works have emerged depending on the kind of ontologies being evaluated and for what purpose [1, 3, 5–7, 9, 15, 20–22]. These works present different quality measures and evaluate some ontologies quantitatively. However, specific studies have not been found about the suitable values of these measures, their acceptance thresholds, and their impact on the quality of the evaluated ontologies.

Quality is not a property of something, but a judgment, so that should be in relation to some purpose [6]. While issues such as orphan classes or consistency in naming are important, the purpose for which the ontology is developed should guide the evaluation of quality thus contributing to the enrichment of its quality. The set of measures and their corresponding weights should be in relation with the purpose of the ontology [15].

A proposed framework to evaluate an ontology considering its specific purpose is OntoQualitas, which includes known measures and new measures to evaluate the quality of an ontology whose purpose is the interchange of information in a collaborative business processes environment [15]. To this aim, a set of requirements is identified that the ontology should fulfill and, associated with them, it is identified a set of questions that reflect specific aspects relevant to the evaluation of ontology. For each question, appropriate measures, their ranges of possible values, and the optimal values are defined. However, the framework does not propose acceptance thresholds of the measure values.

In order to advance in the definition of these thresholds and their impact on the ontology quality, the definition of the proposed measures should be analyzed and, if necessary, modified to ensure their homogeneity. Then, it is necessary to calculate the measures on a set of ontologies and conduct a descriptive statistical analysis of the redefined measures in order to study their behavior.

This paper presents the reformulation of some of the measures outlined in OntoQualitas, resulting in measures that will be more convenient for evaluation of the ontology quality. In addition, a statistical study of a set of ontologies is shown, to whom the reformulated measures were calculated.

The paper is organized as follows: Section 2 describes the main characteristics of the OntoQualitas framework; Section 3 presents the reformulated measures; Section 4 presents the results of the preliminary analysis of data. Results are discussed in Section 5, which also includes the conclusions of this work.

## 2 OntoQualitas

OntoQualitas is a framework to evaluate the quality of an ontology whose purpose is the interchange of information between different contexts [15]. It is structured from an overall requirement imposed on ontologies regarding its content and structure, which is that the ontology should allow the interchange of information between different contexts without imposing a global meaning of such information to all involved contexts. From this overall requirement, three specific requirements are derived: (i) the representation of information interchanged should be formal, (ii) only the information strictly necessary for the interchange must be represented, and (iii) the representation must allow a correct interpretation of the interchanged information in all involved contexts.

The second requirement aforementioned has two aspects: completeness and conciseness. The third requirement has three aspects: semantic correctness, syntactic correctness, and representation correctness, which is assessing the quality of mappings of entities, relations, and features into the elements of the ontology.

OntoQualitas specifies questions that help addressing relevant aspects for ontology evaluation. For each question, appropriate measures are associated. Some of them have been proposed with the objective of assessing the quality of ontologies from a quantitative perspective [3, 5, 6, 20, 21]; others were proposed with the aim of evaluating the mapping between domain entities, its relationships and features, and the elements used for its representation [13].

### 3 Analysis of Measures

In OntoQualitas, the value of some measures is provided in the range  $[0, 1]$ , others are provided in the range  $[0, n]$ , some optimal values are 1, and others are 0. In order to quantify the different quality aspects and to compare values among ontologies, it is necessary to homogenize the value ranges and optimal values of the measures associated with each aspect. As a consequence, a first activity was to modify the definition of some measures to ensure that all have the same scale  $([0, 1])$  and optimal value (1). Additionally, some measures can only be calculated if the considered ontology has the corresponding characteristics. These situations are explicitly identified in Tables 1 to 5.

Completeness (Table 1) refers to the extension, degree, amount or coverage to which the information in a user-independent ontology covers the information of the real world [11].

Concise (Table 2) refers to whether an ontology does not store any unnecessary or useless definitions, if explicit redundancies do not exist between definitions, and redundancies cannot be inferred using other definitions and axioms [11].

Syntactic correctness (Table 3) tries to evaluate the quality of the ontology according to the way it is written, i.e. the correctness and breadth of syntax used [5].

Semantic correctness (Table 4) deals with the vocabulary used to represent entities, relations, and features, and the correctness of the representation of the interchanged information in the ontology.

Representation correctness (Table 5) is related to the quality of mappings of entities, relations, and features into the elements of the ontology evaluated.

### 4 Results of Preliminary Analysis of Data

The results of this preliminary analysis are presented according to the second and third requirements. Since the considered ontologies are formalized in OWL2, the representation of information interchanged is formal, thus achieving the first requirement.

In order to evaluate reformulations to the OntoQualitas measures, ontologies for information interchange between different contexts were needed. A set of ontologies created by students from the course “Development of ontology-based information systems” have been developed from the same specific instructions. First, ontologies (called “base”) were developed by using an ontology learning technique. Then, the representation of entities, their relationships and features were enriched, using a proposed method [14]. These ontologies were called “enriched”. Measures were calculated semi-automatically and the instructions were the frame of reference.

In the base ontologies, certain measures could not be calculated due to lack of the corresponding characteristics. Therefore, in the subsequent statistical analysis, the amount of data varies.

**Table 1.** Completeness measures

Measure	
Necessary and sufficient conditions [11]	$NSC = NSLC/LC$
<i>NSLC</i> : Number of leaf classes with at least one set of necessary and sufficient conditions	
<i>LC</i> : Number of leaf classes	
The ontology should have at least a class hierarchy, without considering the root class ( <b>Thing</b> )	
Existential and universal restrictions [11]	$EUR = EURP/URP$
<i>EURP</i> : Number of properties with existential and universal restrictions along the same property	
<i>URP</i> : Number of properties with universal restrictions	
The ontology should have at least a property with an universal restriction	
Domains and ranges of relations [11]	$DRR = NHRDR/NHR$
<i>NHRDR</i> : Number of non-hierarchical relations with domain and range specified	
<i>NHR</i> : Number of non-hierarchical relations	
The ontology should have at least an object property defined	
* No omission of subclass partition	$NOSP = SPD/CSC$
<i>SPD</i> : Number of subclass-partitions defined on classes with the corresponding disjoint constraint	
<i>CSC</i> : Number of classes with a set of direct subclasses identified	
The ontology should have at least a class hierarchy, without considering the root class ( <b>Thing</b> )	
* No omission of exhaustive subclass partition	$NOESP = CCA/CDSC$
<i>CCA</i> : Number of classes with a set of disjoint direct subclasses and a covering axiom	
<i>CDSC</i> : Number of classes with a set of disjoint direct subclasses identified	
The ontology should have at least a class hierarchy, with a set of disjoint direct subclasses	
Coverage of classes [12]	$Coverage(O_c; F_c) =  O_c \cap F_c  /  F_c $
<i>O<sub>c</sub></i> : Set of classes in the ontology	
<i>F<sub>c</sub></i> : Set of classes in a frame of reference	
The frame of reference should have at least a class	
Coverage of relations between classes [12]	$Coverage(O_{rc}; F_{rc}) =  O_{rc} \cap F_{rc}  /  F_{rc} $
<i>O<sub>rc</sub></i> : Set of relations between classes in the ontology	
<i>F<sub>rc</sub></i> : Set of relations between classes in a frame of reference	
The frame of reference should have at least a relation between classes	
Coverage of relations between instances [12]	$Coverage(O_{ri}; F_{ri}) =  O_{ri} \cap F_{ri}  /  F_{ri} $
<i>O<sub>ri</sub></i> : Set of relations between instances in the ontology	
<i>F<sub>ri</sub></i> : Set of relations between instances in a frame of reference	
The frame of reference should have at least a relation between instances	
Coverage of instances [12]	$Coverage(O_i; F_i) =  O_i \cap F_i  /  F_i $
<i>O<sub>i</sub></i> : Set of instances in the ontology	
<i>F<sub>i</sub></i> : Set of instances in a frame of reference	
The frame of reference should have at least an instance	
Coverage of entity features [15]	$Coverage(O_{fc}; F_{fc}) =  O_{fc} \cap F_{fc}  /  F_{fc} $
<i>O<sub>fc</sub></i> : Set of entity features in the ontology	
<i>F<sub>fc</sub></i> : Set of entity features in a frame of reference	
The frame of reference should have at least an entity feature	
Coverage of dimensions [15]	$Coverage(O_{dfc}; F_{dfc}) =  O_{dfc} \cap F_{dfc}  /  F_{dfc} $
<i>O<sub>dfc</sub></i> : Set of dimensions used to specify entity contextual features in the ontology	
<i>F<sub>dfc</sub></i> : Set of dimensions used to specify entity contextual features in a frame of reference	
The frame of reference should have at least a dimension used to specify entity contextual features	
* The measure was redefined	

**Table 2.** Conciseness measures

Measure	
* Semantically different classes	$SDC = 1 - CSD/C$
<i>CSD</i> : Number of classes with the same formal definition as other class in the ontology	
<i>C</i> : Number of classes in the ontology, without considering the root class ( <b>Thing</b> )	
The ontology should have at least a class hierarchy, without considering the root class ( <b>Thing</b> )	
* Semantically different instances	$SDI = 1 - ISD/I$
<i>ISD</i> : Number of instances with the same formal definition as other instance in the ontology	
<i>I</i> : Number of instances in the ontology	
The ontology should have at least an instance	
* Nonredundant subclass-of relations	$NRSR = 1 - RSCR/HR$
<i>RSCR</i> : Number of redundant subclass-of relations in the ontology	
<i>HR</i> : Number of hierarchical relations	
The ontology should have at least a hierarchical relation, without considering the root class ( <b>Thing</b> )	
* Other nonredundant relations	$ONRR = 1 - RNHR/NHR$
<i>RNHR</i> : Number of redundant non-hierarchical relations in the ontology	
<i>NHR</i> : Number of non-hierarchical relations	
The ontology should have at least a non-hierarchical relation	
* Nonredundant instance-of relations	$NRIR = 1 - RIOR/IOR$
<i>RIOR</i> : Number of redundant instance-of relations in the ontology	
<i>IOR</i> : Number of instance-of relations in the ontology	
The ontology should have at least an instance-of relation	
Precision of classes [12]	$Precision(O_c; F_c) =  O_c \cap F_c  /  O_c $
<i>O<sub>c</sub></i> : Set of classes in the ontology	
<i>F<sub>c</sub></i> : Set of classes in a frame of reference	
The ontology should have at least a class, without considering the root class ( <b>Thing</b> )	
Precision of relations between classes [12]	$Precision(O_{rc}; F_{rc}) =  O_{rc} \cap F_{rc}  /  O_{rc} $
<i>O<sub>rc</sub></i> : Set of relations between classes in the ontology	
<i>F<sub>rc</sub></i> : Set of relations between classes in a frame of reference	
The ontology should have at least a relation between classes	
Precision of entity features [12]	$Coverage(O_{fc}; F_{fc}) =  O_{fc} \cap F_{fc}  /  O_{fc} $
<i>O<sub>fc</sub></i> : Set of entity features in the ontology	
<i>F<sub>fc</sub></i> : Set of entity features in a frame of reference	
The ontology should have at least an entity feature	
Precision of instances [12]	$Precision(O_i; F_i) =  O_i \cap F_i  /  O_i $
<i>O<sub>i</sub></i> : Set of instances in the ontology	
<i>F<sub>i</sub></i> : Set of instances in a frame of reference	
The ontology should have at least an instance	
* The measure was redefined	

**Table 3.** Syntactic correctness measures

Measure	
Lawfulness [5]	$SL = Xb/NS$
<i>Xb</i> : Total breached syntactical rules	
<i>NS</i> : Number of statements in the ontology	
The ontology should have at least a statement	
Richness [5]	$R = Z/Y$
<i>Z</i> : Number of syntactic features used in the ontology	
<i>Y</i> : Number of syntactic features available in the ontology language	
The ontology language should have at least a syntactic feature	

**Table 4.** Semantic correctness measures

Measure	
Interpretability [5]	$IN = SW/WCP$
<i>SW</i> : Number of words used to define classes and properties that have at least a sense listed in WordNet	
<i>WCP</i> : Number of different words used to define classes and properties in the ontology	
The ontology should have at least a class or property name	
* Clarity	$CL = TN / \sum_i S_i$
<i>TN</i> : Total of class or property names in the ontology that have at least a sense listed in WordNet	
<i>S<sub>i</sub></i> : Number of word senses for <i>N<sub>i</sub></i> in WordNet, where <i>N<sub>i</sub></i> is the name of the class or property <i>i</i>	
The ontology should have at least a class or property name that has at least a sense listed in WordNet	
* Non-circularity errors at distance 0	$NCE0 = 1 - Cycles(O; 0)/HR$
<i>Cycles(O; 0)</i> : Number of cycles detected between a class with itself	
<i>HR</i> : Number of hierarchical relations, without considering the root class ( <b>Thing</b> )	
The ontology should have at least a hierarchical relation, without considering the root class ( <b>Thing</b> )	
* Non-circularity errors at distance 1	$NCE1 = 1 - Cycles(O; 1)/HR$
<i>Cycles(O; 1)</i> : Number of cycles detected between a class and an adjacent class	
<i>HR</i> : Number of hierarchical relations, without considering the root class ( <b>Thing</b> )	
The ontology should have at least a hierarchical relation, without considering the root class ( <b>Thing</b> )	
* Non-circularity errors at distance <i>d</i>	$NCEd = 1 - Cycles(O; d)/HR$
<i>Cycles(O; d)</i> : Number of cycles detected between a class and another at <i>d</i> classes away	
<i>HR</i> : Number of hierarchical relations, without considering the root class ( <b>Thing</b> )	
The ontology should have at least a hierarchical relation, without considering the root class ( <b>Thing</b> )	
* Subclass partition without common instances	$SPNCI = 1 - SPCCI/I$
<i>SPCCI</i> : Number of instances that belong to more than one subclass of a partition in the ontology	
<i>I</i> : Number of instances in the ontology	
The ontology should have at least an instance	
* Subclass partition without common classes	$SPNCC = 1 - SPCC/C$
<i>SPCC</i> : Number of classes belonging to more than one subclass of a partition in the ontology	
<i>C</i> : Number of classes in the ontology, without considering the root class ( <b>Thing</b> )	
The ontology should have at least a class, without considering the root class ( <b>Thing</b> )	
* Exhaustive subclass partition without common instances	$ESPNCI = 1 - ESPCCI/I$
<i>ESPCCI</i> : Number of instances belonging to more than one subclass of an exhaustive partition in the ontology	
<i>I</i> : Number of instances in the ontology	
The ontology should have at least an instance	
* Exhaustive subclass partition without common classes	$ESPNC = 1 - ESPCC/C$
<i>ESPCC</i> : Number of classes belonging to more than one subclass of an exhaustive partition in the ontology	
<i>C</i> : Number of classes in the ontology, without considering the root class ( <b>Thing</b> )	
The ontology should have at least a class, without considering the root class ( <b>Thing</b> )	
* Exhaustive subclass partition without external instances	$ESPNEI = 1 - ESPEI/I$
<i>ESPEI</i> : Number of instances of a base class that do not belong to any class of the exhaustive subclass partition of the base class	
<i>I</i> : Number of instances in the ontology	
The ontology should have at least an instance	
* The measure was redefined	

**Table 5.** Representation correctness measures

Measure	
Principle of entities [15] <i>E</i> : number of entities The ontology should have at least an entity	$PE = \sum_k \alpha_k / E$
Principle of intended use of entities [15] <i>U</i> : number of intended uses for all entities The ontology should have at least an intended use for an entity	$PU = \sum_k \alpha_k / U$
Principle of entity relations [15] <i>RE</i> : number of relations identified for all entities The ontology should have at least a relation between entities	$PR = \sum_k \alpha_k / RE$
Principle of simple entity features [15] <i>CS</i> : number of simple entity features identified for all entities The ontology should have at least a simple entity feature	$PCS = \sum_k \alpha_k / CS$
Principle of simple, measurable entity features [15] <i>CM</i> : number of simple, measurable entity features identified for all entities The ontology should have at least a simple, measurable entity feature	$PCM = \sum_k \alpha_k / CM$
Principle of complex entity features [15] <i>CC</i> : number of complex entity features identified for all entities The ontology should have at least a complex entity feature	$PCC = \sum_k \alpha_k / CC$
Principle of common entity features [15] <i>Cc</i> : number of common entity features identified for all entities The ontology should have at least a common entity feature	$PCC = \sum_k \alpha_k / Cc$

$\alpha_k = 0$  if the  $k$  element is not represented;  $\alpha_k = 0.5$  if the  $k$  element is represented in an incomplete form; and  $\alpha_k = 1$  if the  $k$  element is well represented

#### 4.1 Evaluation of Measures

A statistical treatment of the data was performed in order to highlight the most important quality characteristics of ontologies and synthesize them by a few parameters. A total of 39 measures were calculated semi-automatically to a set of 8 ontologies. InfoStat, Student Version<sup>3</sup>, was used to do the statistical analysis of this set of measures. Mean lets see the behavior of each measure on the set of ontologies; position measures, the dispersion of data (deviation; Q1, first quartile; and Q3, third quartile).

**Table 6.** Statistical of completeness measures

Variable	<i>n</i>	Mean	S.D.	Min	Max	Q1	Median	Q3
<i>NSC</i>	6	0,17	0,41	0,00	1,00	0,00	0,00	0,00
<i>EUR</i>	8	0,15	0,35	0,00	1,00	0,00	0,00	0,00
<b><i>DRR</i></b>	<b>8</b>	<b>0,63</b>	<b>0,41</b>	<b>0,00</b>	<b>1,00</b>	<b>0,00</b>	<b>0,82</b>	<b>0,87</b>
<i>NOSP</i>	6	0,09	0,13	0,00	0,30	0,00	0,03	0,19
<i>NOESP</i>	3	0,50	0,50	0,00	1,00		0,50	
<i>Coverage(Oc;Fc)</i>	8	0,22	0,07	0,16	0,36	0,16	0,20	0,24
<i>Coverage(Orc;Frc)</i>	8	0,36	0,23	0,10	0,70	0,10	0,35	0,50
<i>Coverage(Ofc;Ffc)</i>	8	0,02	0,03	0,00	0,07	0,00	0,01	0,04
<b><i>Coverage(Ord;Frd)</i></b>	<b>8</b>	<b>0,93</b>	<b>0,15</b>	<b>0,58</b>	<b>1,00</b>	<b>0,90</b>	<b>1,00</b>	<b>1,00</b>

**Table 7.** Statistical of conciseness measures

Variable	<i>n</i>	Mean	S.D.	Min	Max	Q1	Median	Q3
<b><i>SDC</i></b>	<b>6</b>	<b>0,75</b>	<b>0,28</b>	<b>0,39</b>	<b>1,00</b>	<b>0,53</b>	<b>0,78</b>	<b>1,00</b>
<b><i>SDI</i></b>	<b>4</b>	<b>1,00</b>	<b>0,00</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>
<b><i>NRSR</i></b>	<b>6</b>	<b>1,00</b>	<b>0,00</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>
<b><i>ONRR</i></b>	<b>7</b>	<b>0,86</b>	<b>0,38</b>	<b>0,00</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>
<b><i>NRIR</i></b>	<b>4</b>	<b>1,00</b>	<b>0,00</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>	<b>1,00</b>
<i>Precision(Oc;Fc)</i>	8	0,33	0,20	0,07	0,57	0,08	0,31	0,50
<i>Precision(Orc;Frc)</i>	8	0,16	0,16	0,02	0,38	0,03	0,07	0,33
<i>Precision(Ofc;Ffc)</i>	8	0,00	0,00	0,00	0,00	0,00	0,00	0,00
<i>Precision(Oi;Fi)</i>	4	0,00	0,00	0,00	0,00	0,00	0,00	0,00

<sup>3</sup> <http://www.infostat.com.ar/>

Regarding completeness (Table 6), only two of the nine measures have a mean greater than 0.6. The measure with the highest mean value is *Coverage of dimensions* ( $Coverage(O_{dfc}; F_{dfc})$ ); 90.0% of ontologies have a value greater than or equal to 0.9, meaning that most of the dimensions used to specify entity contextual features were made explicit in the ontology. Then, *Domains and ranges of relations* (*DRR*) follows with a mean of 0.63, which determines the proportion of domain and range of the relations and functions exactly and precisely delimited. The frame of reference had no instances. Then, the measures *Coverage of relations between instances* and *Coverage of instances*, not listed in Table 6, could not be calculated.

In regards to conciseness (Table 7), except in *Precision*, all other measures have high values. Half of ontologies have all of instances semantically different and nonredundant instance-of relations (*SDI* and *NRIR* are optimal). The other half has no instances. No ontologies with hierarchical relations have redundant subclass-of relations (*NRSR* has optimum value in all measures). *Semantically different classes* (*SDC*) has a mean of 0.75 and 75% of ontologies have a value greater than or equal to 0.53, meaning that more than half of subclasses are defined with different characteristics. 75% of ontologies do not have redundant non-hierarchical relations (*ONRR* is optimal).

**Table 8.** Statistical of semantic correctness measures

Variable	<i>n</i>	Mean	S.D.	Min	Max	Q1	Median	Q3
<i>IN</i>	8	<b>0,52</b>	<b>0,28</b>	<b>0,21</b>	<b>0,96</b>	<b>0,31</b>	<b>0,43</b>	<b>0,50</b>
<i>CL</i>	8	<b>0,34</b>	<b>0,14</b>	<b>0,13</b>	<b>0,52</b>	<b>0,15</b>	<b>0,40</b>	<b>0,40</b>
<i>NCE0</i>	6	1,00	0,00	1,00	1,00	1,00	1,00	1,00
<i>NCE1</i>	6	1,00	0,00	1,00	1,00	1,00	1,00	1,00
<i>NCEd</i>	6	1,00	0,00	1,00	1,00	1,00	1,00	1,00
<i>SPNCI</i>	4	0,75	0,50	0,00	1,00	0,00	1,00	1,00
<i>SPNCC</i>	8	0,96	0,06	0,88	1,00	0,88	1,00	1,00
<i>ESPNCI</i>	4	1,00	0,00	1,00	1,00	1,00	1,00	1,00
<i>ESPNCC</i>	8	1,00	0,00	1,00	1,00	1,00	1,00	1,00
<i>ESPNEI</i>	4	1,00	0,00	1,00	1,00	1,00	1,00	1,00

**Table 9.** Statistical of syntactic correctness measures

Variable	<i>n</i>	Mean	S.D.	Min	Max	Q1	Median	Q3
<i>SL</i>	8	1,00	0,00	1,00	1,00	1,00	1,00	1,00
<b><i>R</i></b>	<b>8</b>	<b>0,05</b>	<b>0,03</b>	<b>0,01</b>	<b>0,11</b>	<b>0,03</b>	<b>0,03</b>	<b>0,05</b>

**Table 10.** Statistical of representation correctness measures

Variable	<i>n</i>	Mean	S.D.	Min	Max	Q1	Median	Q3
<b><i>PE</i></b>	<b>8</b>	<b>0,10</b>	<b>0,04</b>	<b>0,06</b>	<b>0,17</b>	<b>0,06</b>	<b>0,11</b>	<b>0,13</b>
<b><i>PU</i></b>	<b>8</b>	<b>0,90</b>	<b>0,09</b>	<b>0,80</b>	<b>1,00</b>	<b>0,83</b>	<b>0,88</b>	<b>1,00</b>
<i>PR</i>	8	0,36	0,23	0,10	0,70	0,10	0,35	0,50
<b><i>PCS</i></b>	<b>4</b>	<b>0,90</b>	<b>0,13</b>	<b>0,75</b>	<b>1,00</b>	<b>0,75</b>	<b>0,92</b>	<b>1,00</b>
<i>PCM</i>	4	0,63	0,48	0,00	1,00	0,00	0,75	1,00

In relation to the semantic correctness (Table 8), the measures are mostly high. The hierarchies are well defined, without cycles (*NCE0*, *NCE1*, and *NCEd*), as well as the exhaustive subclass partitions (*ESPNCI*, *ESPNCC*, and *ESPNEI*). By contrast, ontologies are moderately interpretable and unclear; 75% of them have a value less than or equal to 0.5 and 0.4, respectively.

As for syntactic correctness (Table 9), it can be observed that the ontologies are syntactically correct, but the proportion of syntactic features used is very low, despite the development of ontologies supported by a case tool.

Finally, as to the representation correctness (Table 10), on average, 90% of the intended use and simple features of entities is represented according to its principle. However, only in 10% of cases, on average, the representation of entities is performed through classes of ontology. The measures *Principle of*



*complex entity features* and *Principle of common entity features* could not be calculated because the ontologies do not have these characteristics.

## 5 Discussion and Conclusions

In this paper, the reformulation of some measures of the OntoQualitas framework has been presented, and the results of a preliminary analysis over the values obtained from applying such measures to a set of ontologies have been shown.

According to the results, the evaluated ontologies do not fulfill adequately the second requirement, i.e., the representation of the information strictly necessary for the interchange. In part, this may be due to the ontology learning tool used to generate the base ontologies that do not add necessary and sufficient conditions, or existential and universal restrictions, among others.

Looking at the syntactic correctness measures, it can be observed that the richness of language was not seized, despite the use of case tools for the development of ontologies. The use of ontology learning techniques contributed to this, as only limited to map the elements of the source into the ontology language elements, untapped all syntactic features available.

As for the semantic interpretation, measures revealed that the names for the ontology elements (classes, relations, properties) were not properly selected.

Regarding the representation correctness, an unexpected result is the low representation of entities through the ontology classes.

Finally, these measures allow detecting errors in the development of ontologies, which affects its quality. An exploratory analysis of the data allowed to characterize the studied ontologies. Future work is to carry out an inferential statistical analysis to a larger set of ontologies that allows analyzing the possible interdependence between measures, define acceptance thresholds of measures, and propose a strategy for assessing the quality of ontologies.

## References

1. Alani, H., Brewster, C., Shadbolt, N.: Ranking ontologies with AKTiveRank. 5th International Semantic Web Conference, ISWC. LNCS 4273, 1–15 (2006)
2. Álvarez Suárez, M.M., Caballero, A., Pérez Lechuga, G.: Análisis multivariante: Clasificación, organización y validación de resultados. 4th Int. Latin American & Caribbean Conference for Engineering and Technology (LACCET) (2006)
3. Brank, J., Grobelink, M., Mladenic, D.: A survey of ontology evaluation techniques. Conference on Data Mining and Data Warehouses (SiKDD), 166–169 (2005)
4. Breitman, K.K., Casanova, M.A., Truskowski, W.: Semantic Web: Concepts, technologies and applications. NASA monographs in systems and software engineering. Springer-Verlag London Limited (2007)
5. Burton-Jones, A., Storey, V.C., Sugumaran, V., Ahluwalia, P.: A semiotic metrics suite for assessing the quality of ontologies. Data Knowl. Eng. 55(1), 84–102 (2005)
6. Colomb, R.M.: Quality of ontologies in interoperating information systems. Technical report 18/02 ISIB-CNR, Padova, Italy (2002)

7. Duque-Ramos, A., Fernández-Breis, J., Stevens, R., Aussenac-Gilles, N.: OQuaRE: A SQuaRE-based approach for evaluating the quality of ontologies. *J. Res. Pract. Inf. Tech.* 43(2), 159–176 (2011)
8. Duque-Ramos, A., Fernández-Breis, J., Iniesta, M., Dumontier, M., Egaña Aranguren, M., Schulz, S., Aussenac-Gilles, N., Stevens, R.: Evaluation of the QuaRE framework for ontology quality. *Expert Syst. Appl.* 40, 2669–2703 (2013)
9. Gangemi, A., Catenacci, C., Ciaramita, M., Lehmann, J.: Ontology evaluation and validation. *The Semantic Web: Research and Applications. 3rd European Semantic Web Conference, ESWC, Proceedings, LNCS 4011*, 140–154 (2006)
10. Gašević, D., Djurić, D., Devedžić, V.: Model driven architecture and ontology development. Springer-Verlag New York, Inc., Secaucus, NJ, USA (2006)
11. Gómez-Pérez, A.: Evaluation of ontologies. *Int. J. Intell. Syst.* 16, 391–409 (2001)
12. Guarino, N.: Towards a formal evaluation of ontology quality. *IEEE Intell. Syst.* 19(4), 74–81 (2004)
13. Rico, M.: Soporte para enriquecer la representación de entidades en una ontología. Tesis doctoral, Universidad Tecnológica Nacional, Fac. Reg. Santa Fe, AR (2011)
14. Rico, M., Caliusco, M.L., Chiotti, O., Galli, M.R.: An approach to define semantics for BPM systems interoperability. *Enterprise Information Systems*, DOI:10.1080/17517575.2013.767381 (2013)
15. Rico, M., Caliusco, M.L., Chiotti, O., Galli, M.R.: OntoQualitas: A framework for ontology quality assessment in information interchanges between heterogeneous systems. *Comput. Ind.* 65(9), 1291–1300 (2014)
16. Romero Villafranca, R.: Curso de introducción a los métodos de análisis estadístico multivariante. Universitat Politècnica de València, SP.UPV.95–606 (1995)
17. Simperl, E., Mochol, M., Bürger, T.: Achieving maturity: The state of practice in ontology engineering in 2009. *Int. J. Comput. Sci. Appl.* 7(1), 45–65 (2010)
18. Staab, S., Studer, R. (Eds.): *Handbook on ontologies. International handbooks on information systems. 2nd edn.* Springer-Verlag Berlin Heidelberg (2009)
19. Studer, R., Benjamins, V.R., Fensel, D.: Knowledge engineering: Principles and methods. *Data. Knowl. Eng.* 25(1-2), 161–197 (1998)
20. Stvilia, B.: A model for ontology quality evaluation. *First Monday* 12(12) (2007)
21. Tartir, S., Arpinar, I.B.: Ontology evaluation and ranking using OntoQA. *International Conference on Semantic Computing, ICSC*, 185–192 (2007)
22. Vrandečić, D.: Ontology evaluation. PhD Thesis. Institute AIFB, University of Karlsruhe, Germany (2010)