



# OntoQualitas: A framework for ontology quality assessment in information interchanges between heterogeneous systems



Mariela Rico <sup>a,\*</sup>, María Laura Caliusco <sup>a</sup>, Omar Chiotti <sup>a,b</sup>, María Rosa Galli <sup>a,b</sup>

<sup>a</sup> CIDISI Research Center, UTN-FRSF, Lavalse 610, S3004EWB Santa Fe, Argentina

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

## ARTICLE INFO

### Article history:

Received 12 July 2013

Received in revised form 13 June 2014

Accepted 28 July 2014

Available online 5 September 2014

### Keywords:

Ontology evaluation

Information interchange

Measurement

Ontology requirement

## ABSTRACT

Nowadays, Internet technologies and standards are being systematically used by enterprises as tools to provide an infrastructure to connect people, enterprises, and applications they are using. In such complex networked enterprises, it is increasingly challenging to interchange, share, and manage internal and external digital information. In this context, to achieve interoperability between information systems is a challenging task. In order to solve the interoperability problem at semantic level, several ontology-based approaches have emerged. Although methodologies, methods, techniques, and tools to support the ontology building process were proposed, there are no mature models to measure this process, and the quality of implemented ontologies remains a major concern. This paper presents a framework, OntoQualitas, for evaluating the quality of an ontology whose purpose is the information interchange between different contexts. OntoQualitas includes previous and new measures to evaluate the ontology considering its specific purpose. Additionally, an empirical validation of OntoQualitas is presented.

© 2014 Elsevier B.V. All rights reserved.

## 1. Introduction

In the new competitive and dynamic global market, industrial enterprises are establishing partnerships by creating new business models and are using Internet technologies and standards as tools to provide an infrastructure to connect their information systems. In such complex emerging networked enterprises, to achieve the interoperability between information systems is a challenging task. Three levels of interoperability have been highlighted: technical, semantic, and organizational [1]. In order to solve the interoperability problem at semantic level, several approaches based on the use of ontologies have been proposed [2–5].

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 [6]. In order to assess the ontology quality, different works have emerged depending on the kind of ontologies being evaluated and for what purpose [7–12]. Regarding the purpose, some approaches intend to rank ontologies and others offer information about the strengths and weaknesses of the

ontologies. The present work goes in the latter direction, similar to approach OQuaRE [13].

OQuaRE is a framework for ontology quality assessment based on the SQuaRE standard for software quality evaluation [13]. To this aim, ontologies are considered as software artifacts. The framework includes the quality model and the quality measures, although the usage of SQuaRE requires also the definition of quality requirements and quality evaluation. The quality model is comprised of a set of quality characteristics: reliability, operability, maintainability, compatibility, transferability, and functional adequacy. Structural features of ontologies, not considered in the standard, were added to the framework. Each characteristic has a series of subcharacteristics associated. The quality subcharacteristics have measures associated; some of them were adapted from Object-Oriented Programming measures and others from those developed by the ontology engineering community, especially for structural properties.

While OQuaRE can be seen as a step toward standardization of the ontology evaluation process, an important issue is not considered. An ontology whose purpose is the interchange of information between different contexts represents semantics of information interchanged. Quality is not a property of something, but a judgment, so that should be in relation to some purpose [9]. Trying to evaluate the quality of an ontology without considering its intended uses is meaningless. The set of characteristics,

\* Corresponding author. Tel.: +54 342 4602390.

E-mail address: [mrlico@santafe-conicet.gov.ar](mailto:mrlico@santafe-conicet.gov.ar) (M. Rico).

**Table 1**  
OntoQualitas: role and requirements of an ontology.

<i>Role</i>	The ontology is used to represent the semantics of information interchanged between heterogeneous information systems that belong to different contexts.
<i>Requirements</i>	The overall requirement with regards to the content and structure of the ontology is to allow the interchange of information between different contexts without imposing a global meaning of such information to all the involved contexts. From this requirement, three specific requirements (OR) were derived: OR1: Formal representation of the interchanged information. OR2: Representation of the information strictly necessary for the interchange. OR3: Correct interpretation of the information interchanged in all contexts involved.

sub-characteristics, measures, and the corresponding weights should be in relation with the purpose of the ontology.

This paper presents the OntoQualitas framework for evaluating the quality of an ontology whose purpose is the interchange of information between different contexts. OntoQualitas includes previous measures and new measures proposed to evaluate the ontology considering its specific purpose.

The paper is organized as follows: Section 2 discusses the ontology properties for information interchange. Section 3 presents a framework for ontology quality assessment in information interchange. Section 4 presents an empirical validation of the framework. Finally, Section 5 is devoted to conclusions.

## 2. Ontology for information interchange

The evolution of information systems and communication technologies has been leading to the implementation of Peer-to-Peer communication models between heterogeneous information systems [14]. This communication model allows maintaining some forms of peer privacy and autonomy, whence, they may use different terminologies and metadata models even if they refer to the same domain [15]. Thus, in order to interchange information between information systems, ontologies are used to define the semantics of the information to be shared.

Two ontologies belonging to the same domain could present differences in their naming conventions, structure, and/or in the representation way [16]. This implies a new challenge: how to allow the interoperability between heterogeneous ontologies, each one representing an information source, to guarantee semantic interoperability between information systems that interchange information between these sources.

Ontology-matching is a plausible solution to allow the interoperability between heterogeneous ontologies [15–17]. It aims at finding correspondences between semantically related entities of different ontologies. However, there are certain conflicts that ontology-matching techniques are unable to solve such as how entities represented in ontologies are interpreted. For example, if the entity Product appears in an ontology of a packaging industry, it should not be concluded that it is equivalent to the entity Product in an ontology of a dairy industry. Associated with each of these entities is an intended use of the entity, generally missing in its representation. When ontologies are developed without this consideration, it is impossible for the matching process to detect if those entities are equivalent or not. In these cases, conflicts are related to incomplete representations in ontologies to be matched. Then, in information systems interchanging information the representation of the real meaning of shared entities is crucial, and quality evaluations of the implemented ontologies are needed.

The quality of an ontology whose purpose is the interchange of information between heterogeneous systems can be measured by its capabilities to (i) provide a formal representation of the interchanged information, and (ii) meet information systems requirements.

In this work, ontologies formalized in OWL 2 [18] are considered for the purpose of describing the evaluation measures. The term class refers to both primitive class and defined class, except where indicated otherwise.

## 3. A framework for ontology quality assessment in information interchange

Ontology evaluation methodologies provide a framework for defining appropriate methods for evaluating ontologies [19]. OntoClean [20], OntoMetric [21], and ROMEO [19] are the most recognized. The OntoClean methodology focuses on evaluating the subsumption relation between classes in an ontology, whereas OntoMetric methodology focuses on selecting an ontology based on a manual specification of the importance of ontology aspects. Different from them, the ROMEO methodology allows evaluating many aspects of an ontology based on the purpose for which that ontology has been designed.

OntoQualitas, built following the ROMEO methodology, allows evaluating an ontology whose purpose is to support the information interchange between different contexts. Initially, the ontology roles, which specify the way the ontology is used, and the set of ontology requirements, which specify the ontology qualities, were defined.

In order to attend the ontology role, in this paper the requirements defined for data integration [22] were specialized. In addition, other requirements, based on several previous studies [3,23–25], were incorporated with the aim of making explicit all the needed aspects to avoid ambiguity and limit the ontology contents to the purpose.

Table 1 shows the role and requirements defined in OntoQualitas.

For each ontology requirement identified, one or more questions were specified. A question reflects a specific quality or aspect relevant for ontology evaluation. For each question, appropriate quantifiable measures, able to sufficiently answer it, were selected.

### 3.1. Evaluating OR1: formal representation of the interchange information

Questions (Q) specified to evaluate how the information interchanged is formally represented are:

Q1.1: Is the language in which the ontology is implemented machine-readable?

One issue that a peer has to face is how to access the information interchanged with other peers [26]. Since this interchange is made between heterogeneous information systems, the ontologies used to define the semantics of the information interchanged must be implemented in a machine executable and machine interpretable language.

Q1.2: Does the language allow the representation of entities, their relations and features?

**Table 2**  
Measures for questions formulated for OR1.

Question	Measure	Value	Optimal
Q1.1	Machine-readable language	0: no – 1: yes	1
Q1.2	Entities definition capability	0: no – 1: yes	1
	Relations definition capability	0: no – 1: yes	1
	Features definition capability	0: no – 1: yes	1
Q1.3	Potential reasoning	0: no – 1: yes	1

The purpose of this question is to answer if the language has elements that permit to adequately represent the semantics of the information interchanged.

Q1.3: Does the language allow making inferences?

This question allows knowing if useful inferences can be drawn with the language.

Table 2 shows the measures selected for evaluating the questions formulated for OR1.

### 3.2. Evaluating OR2: representation of the information strictly necessary

Questions specified to evaluate if the information represented is strictly necessary are:

Q2.1: Does the ontology contain all the information necessary for the interchange?

This question helps to explore the completeness aspect of requirement OR2. *Completeness* refers to the extension, degree, amount or coverage to which the information in a user-independent ontology covers the information of the real world [27]. Completeness depends on the level of granularity agreed. An ontology is complete if and only if all that is supposed to be in the ontology is explicitly set out in it, or can be inferred from axioms.

Question Q2.1 was subdivided as follows:

Q2.1.1: Is each leaf class defined by necessary and sufficient conditions?

A class that only has necessary conditions is known as a primitive class. A class that has at least one set of necessary and sufficient conditions is known as a defined class [28]. Primitive classes are usually found near the top of a generalization hierarchy and defined classes typically appear when descend into a hierarchy by specializing general classes with various restrictions.

Additionally, when describing a class, a universal restriction along a given property should be used with the corresponding existential restriction along the same property.

Q2.1.2: Are the domain and range of the relations and functions exactly and precisely delimited?

Q2.1.3: Is there subclass-partitions defined on classes without the corresponding disjoint constraint over the subclasses sets?

Q2.1.4: Is there disjoint subclass-partitions defined on classes without the corresponding covering axioms over the covered classes?

Q2.1.5: Does the ontology have entities, relations, or instances missing with regards to the information to be interchanged?

Q2.1.6: Does the ontology have features of an entity missing that may be inferred in wrong way in other contexts different from the considered context?

**Table 3**  
Completeness measures.

Question	Measure	Range	Optimal
Q2.1.1	Necessary and sufficient conditions	[0;1]	1
	$NSC = NSLC/LC$ [27]		
Q2.1.2	Existential and universal restrictions	[0;1]	1
	$EUR = EURP/URP$		
Q2.1.2	Domains and ranges of relations	[0;1]	1
	$DRR = NHRDR/NHR$ [27]		
Q2.1.3	Subclass partition omission	[0;1]	0
	$SPO = SPND/CSC$ [27]		
Q2.1.4	Exhaustive subclass partition omission $ESPO$	[0;1]	0
	If $CDSC > 0$ then $ESPO = CNCA/CDSC$ else $ESPO = 0$ [27]		
Q2.1.5	$Coverage(O_c; F_c) =  O_c \cap F_c  /  F_c $ [29]	[0;1]	1
	$Coverage(O_{rc}; F_{rc}) =  O_{rc} \cap F_{rc}  /  F_{rc} $	[0;1]	1
	$Coverage(O_{ri}; F_{ri}) =  O_{ri} \cap F_{ri}  /  F_{ri} $	[0;1]	1
	$Coverage(O_i; F_i) =  O_i \cap F_i  /  F_i $	[0;1]	1
Q2.1.6	$Coverage(O_{fc}; F_{fc}) =  O_{fc} \cap F_{fc}  /  F_{fc} $	[0;1]	1
Q2.1.7	$Coverage(O_{dfc}; F_{dfc}) =  O_{dfc} \cap F_{dfc}  /  F_{dfc} $	[0;1]	1

CDSC: number of classes with a set of disjoint direct subclasses identified.  
 CNCA: number of classes with a set of disjoint direct subclasses and without a covering axiom.  
 CSC: number of classes with a set of direct subclasses identified.  
 EURP: number of properties with existential and universal restrictions along the same property.  
 F<sub>c</sub>: set of classes in a frame of reference.  
 F<sub>rc</sub>: set of relations between entities in a frame of reference.  
 F<sub>ri</sub>: set of relations between instances in a frame of reference.  
 F<sub>i</sub>: set of instances in a frame of reference.  
 F<sub>fc</sub>: set of entity features in a frame of reference.  
 F<sub>dfc</sub>: set of dimensions used to specify entity contextual features in a frame of reference.  
 LC: number of leaf classes.  
 NHR: number of non-hierarchical relations.  
 NHRDR: non-hierarchical relations with domain and range specified.  
 NSLC: leaf classes with at least one set of necessary and sufficient conditions.  
 O<sub>c</sub>: set of classes in the ontology.  
 O<sub>rc</sub>: set of relations between entities in the ontology.  
 O<sub>ri</sub>: set of relations between instances in the ontology.  
 O<sub>i</sub>: set of instances in the ontology.  
 O<sub>fc</sub>: set of entity features in the ontology.  
 O<sub>dfc</sub>: set of dimensions used to specify entity contextual features in the ontology.  
 SPND: number of subclass-partitions defined on classes without the corresponding disjoint constraint.  
 URP: number of properties with universal restrictions.

Q2.1.7: Are the dimensions used to specify contextual features the same regardless of the context in which the features are considered? Are they represented in the ontology?

Table 3 shows the measures selected for evaluating questions Q2.1 formulated for OR2.

Q2.2: Does the ontology contain only the information necessary for the interchange?

This question helps to explore the *conciseness* aspect of requirement OR2. An ontology is concise if it 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 [27]. A concise ontology maps to an ontology that has high precision regarding to the information to be interchanged, that is, it does not model entities, instances, and relations outside of the information to be interchanged [19]. However, conciseness doesn't imply absence of redundancies. In information interchange between heterogeneous information systems that belong to different contexts, some degree of controlled redundancy is necessary in order to represent the possible uses of entities, since these uses depend on the context in which the entities are considered.

**Table 4**  
Conciseness measures.

Question	Measure	Range	Optimal
Q2.2.1	Semantically identical classes $SIC = CSD/C$ [27]	[0;1]	0
	Semantically identical instances $SII = ISD/I$ [27]	[0;1]	0
Q2.2.2	Redundant subclass-of relations $RSR = RSCR/HR$ [27]	[0;1]	0
	Other redundant relations $ORR = RNHR/NHR$	[0;1]	0
	Redundant instance-of relations $RIR = RIOR/IOR$ [27]	[0;1]	0
Q2.2.3	$Precision(O_c;F_c) = (O_c \cap F_c)/O_c$ [29]	[0;1]	1
	$Precision(O_r;F_r) = (O_r \cap F_r)/O_r$	[0;1]	1
	$Precision(O_i;F_i) = (O_i \cap F_i)/O_i$	[0;1]	1
	$Precision(O_i;F_i) = (O_i \cap F_i)/O_i$	[0;1]	1

C: number of classes in the ontology.

CSD: number of classes with the same formal definition as other class in the ontology.

$F_c$ : set of classes in a frame of reference.

$F_r$ : set of relations between classes in a frame of reference.

$F_i$ : set of class features in a frame of reference.

$F_j$ : set of instances in a frame of reference.

HR: number of hierarchical relations.

I: number of instances in the ontology.

IOR: number of instance-of relations in the ontology.

ISD: number of instances with the same formal definition as other instance in the ontology.

NHR: number of non-hierarchical relations.

$O_c$ : set of classes in the ontology.

$O_r$ : set of relations between entities in the ontology.

$O_i$ : set of entity features in the ontology.

$O_j$ : set of instances in the ontology.

RIOR: number of redundant instance-of relations in the ontology.

RNHR: number of redundant non-hierarchical relations in the ontology.

RSCR: number of redundant subclass-of relations in the ontology.

Question Q2.2 was subdivided as follows:

Q2.2.1: How many identical entities or instances are modeled using different names?

Q2.2.2 How many redundant relations are found in the ontology?

Q2.2.3 Does the ontology model entities, relations, features, or instances outside of the needed for the information interchange?

Table 4 shows the measures selected for evaluating questions Q2.2 formulated for OR2.

### 3.3. Evaluating OR3: correct interpretation of the information interchanged in all contexts involved

From the OR3 requirement, five main design principles were proposed for enriching the representation of the interchanged information and avoiding misunderstandings [25]. These design principles (DP) and ontology requirements (OR) derived from them are:

- DP1. An entity must be represented by a class. In this case, only universals are considered as entity, i.e., sets that contain individuals.  
OR3.1: Ensuring the representation of a universal entity is made by a class.
- DP2. Each intended use of an entity must be represented by a class, called bridge, which allows linking meanings and representations of an entity in different contexts. This allows representing the possible uses of entities that depend on the context in which entities are considered, thus avoiding confounding conflicts [22].  
OR3.2: Ensuring the representation of each intended use of an entity is made by a bridge class.

DP3. A relation between entities must be represented by an object property between the classes representing these entities.

OR3.3: Ensuring the representation of a relation between entities is made by an object property between the classes representing these entities.

DP4. An entity feature that can be considered an entity in itself must be represented by a class, making explicit the context of its value and value scaling where appropriate [22]. There are three cases.

DP4.1: A simple entity feature is a quality that does not bear other qualities, and it is associated with a one-dimensional representation [30]. Thus, a simple feature must be represented by:

- A class representing the dimension (the set of possible values).
- An object property between this class and that representing the simple feature.

OR3.4: Ensuring the representation of a simple entity feature is correct.

DP4.2: A simple, measurable entity feature has an associated measurement unit that affects its dimension granularity but not its structure. In turn, measurement unit is associated with a physical dimension. Thus, a simple, measurable feature must be represented by:

- A class representing the dimension.
- An object property between this class and that representing the simple feature.
- A class representing the measurement unit of the dimension.
- An object property between this class and that representing the dimension.
- A class representing the physical dimension of the measurement unit.
- An object property between this class and that representing the measurement unit.

OR3.5: Ensuring the representation of a simple, measurable entity feature is correct.

DP4.3: A complex entity feature is a quality that bears other qualities, and it is associated with a set of integral dimensions that are separable from all other dimensions [30]. An integral dimension is one in which for assigning a value to an object on one dimension, it is necessary giving it a value on other dimensions. Thus, a complex feature must be represented by:

- A class representing the set of integral dimensions.
- An object property between this class and that representing the entity.
- For each integral dimension, a class representing it and an object property between it and that representing the set of integral dimensions.
- For each class representing a measurable, integral dimension, there must be: a class representing the measurement unit of the integral dimension; an object property between this class and that representing the measurable integral dimension; a class representing the physical dimension of the measurement unit; and an object property between this class and that representing the measurement unit.

Additionally, class expressions of existence, cardinality, and closure must be added to ensure that the set of integral dimensions is made by the corresponding integral

**Table 5**  
Syntactic correctness measures.

Question	Measure	Range	Optimal
Q3.1.1	Lawfulness $SL = Xb/NS$ [8]	[0;1]	1
Q3.1.2	Richness $R = Z/Y$ [8]	[0;1]	1

NS: number of statements in the ontology.  
 Xb: total breached syntactical rules.  
 Y: number of syntactic features available in the ontology language.  
 Z: number of syntactic features used in the ontology.

dimensions. Another class axiom to be defined is one requiring that the integral dimensions must be disjoint.  
 OR3.6: Ensuring the representation of a complex entity feature is correct.

DP5. An entity feature that cannot be considered an entity in itself must be represented by a data property.  
 OR3.7: Ensuring the representation of an entity feature that can not be considered an entity in itself is made by a data property.

Questions specified to evaluate if the information interchanged is correctly interpreted in all contexts involved (OR3) are:

Q3.1: Is the ontology syntactically correct? Can a computer read the ontology?

Answer to this question tries to evaluate the quality of the ontology according to the way it is written. When the ontology development is supported by a case tool, the syntactic correctness is richness and complexity of syntax rather than correctness, since the case tool should be able to detect syntactically incorrect structure, wrong keywords in definitions, loops between definitions, and so on. However, when the ontology editor doesn't have error-checking capabilities, lawfulness should be considered [8].

Question Q3.1 was subdivided as follows:

Q3.1.1: Have the language's rules been complied?  
 Q3.1.2: What proportion of features in the ontology language has been used?

Table 5 shows the measures selected for evaluating questions Q3.1 formulated for OR3.

Q3.2: Can a computer understand the ontology?  
 This question deals with the vocabulary used to represent entities, relations, and features, and the correctness of the representation of the interchanged information in the ontology.

Question Q3.2 was subdivided as follows:

Q3.2.1: Will the terms used in the ontology have meaning?  
 The objective of this question is to state the interpretability of terms. The knowledge provided by the ontology should map into meaningful concepts in the real world [8].

Q3.2.2: Is the context of terms used in the ontology clear?  
 This question refers to a criterion of having no ambiguous terms used in an ontology.

Q3.2.3: Is the ontology consistent?

**Table 6**  
Semantic correctness measures.

Question	Measure	Range	Optimal
Q3.2.1	Interpretability $I = SW/WCP$ [8]	[0;1]	1
Q3.2.2	Clarity $CL = \sum_i S_i/TN$ [8]	[0;n]	1
Q3.2.3	Circularity errors at distance 0 = $Cycles(O;0)$ [27]	[0;n]	0
	Circularity errors at distance 1 = $Cycles(O;1)$	[0;n]	0
	Circularity errors at distance $d = Cycles(O;d)$	[0;n]	0
	Subclass partition with common instances [27]	[0;n]	0
	Subclass partition with common classes [27]	[0;n]	0
	Exhaustive subclass partition with common instances [27]	[0;n]	0
	Exhaustive subclass partition with common classes [27]	[0;n]	0
	Exhaustive subclass partition with external instances [27]	[0;n]	0

$Cycles(O;0)$ : number of cycles detected between a class with itself.  
 $Cycles(O;1)$ : number of cycles detected between a class and an adjacent class.  
 $Cycles(O;d)$ : number of cycles detected between a class and another at d classes away.  
 $S_i$ : number of word senses for  $N_i$  in WordNet [31], where  $N_i$  is the name of the class  $i$ .  
 SW: number of words used to define classes and properties that have at least a sense listed in WordNet.  
 TN: total of class or property names in the ontology.  
 WCP: number of different words used to define classes and properties in the ontology.  
 $n \geq 1$ .

Consistency refers to whether it is possible to obtain contradictory conclusions from valid input definitions. A given definition is consistent if and only if the individual definition is consistent and no contradictory sentences can be inferred using other definitions and axioms [27].

Table 6 shows the measures selected for evaluating questions Q3.2 formulated for OR3.

The correctness of the representation of the interchanged information has to do with satisfying requirements OR3.1 to OR3.7. To this aim, questions to evaluate the quality of mappings of entities, relations, and features into the elements of the ontology were specified.

Q3.2.4: Does the ontology include a class for the representation of each universal entity?

Q3.2.5: Does the ontology include a bridge class for the representation of each intended use of an entity?

Q3.2.6: Does the ontology include an object property for the representation of each relation between entities?

Q3.2.7: Does the ontology include all the elements required for the representation of a simple entity feature?

Q3.2.8: Does the ontology include all the elements required for the representation of a simple, measurable entity feature?

Q3.2.9: Does the ontology include all the elements required for the representation of a complex entity feature?

Q3.2.10: Does the ontology include all the elements required for the representation of an entity feature that can not be considered an entity in itself?

Considering a correct representation the one that meet the corresponding principle, and a weight vector, it is proposed to calculate the indicators for the aforementioned principles as defined in Table 7. The  $\alpha$  values were defined as:  $\alpha_k = 0$  if the  $k$

**Table 7**  
Representation correctness measures.

Question	Measure	Range	Optimal
Q3.2.4	Entities $PE = \sum_k \alpha_k/E$	[0;1]	1
Q3.2.5	Intended use of entities $PU = \sum_k \alpha_k/U$	[0;1]	1
Q3.2.6	Entity relations $PR = \sum_k \alpha_k/RE$	[0;1]	1
Q3.2.7	Simple entity features $PCS = \sum_k \alpha_k/CS$	[0;1]	1
Q3.2.8	Simple, measurable entity features $PCM = \sum_k \alpha_k/CM$	[0;1]	1
Q3.2.9	Complex entity features $PCC = \sum_k \alpha_k/CC$	[0;1]	1
Q3.2.10	Common entity features $PCCc = \sum_k \alpha_k/Cc$	[0;1]	1

CC: number of complex entity features identified for all entities.  
 Cc: number of common entity features identified for all entities.  
 CM: number of simple, measurable entity features identified for all entities.  
 CS: number of simple entity features identified for all entities.  
 E: number of entities.  
 RE: number of relations identified for all entities.  
 U: number of intended uses for all entities.

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.

**Q3.3: Is the ontology useful?**

This question relates to the ontology usefulness for heterogeneous information systems involved in the information interchange, regardless of syntax or semantics. Each system has to apply an ontology-matching process between the ontology used to the information interchange and its own ontologies. The found correspondences between semantically related entities of different ontologies should be assessed in order to respond on the usefulness of the ontology used for the interchange.

Table 8 shows the measures selected for evaluating question Q3.3 formulated for OR3.

**4. Empirical study of the ontology measures for information interchange**

This section presents results of an empirical study in order to determinate: suitability of measures to questions, how questions influence ontology qualities, and how ontology qualities influence the quality of an ontology for information interchange between information systems belonging to different contexts. The case study is based on a collaborative relationship between a packaging

**Table 8**  
Usefulness measures.

Question	Measure	Range	Optimal
Q3.3	Expected improvement index	[0;n]	$n$
	$IE = (Ee - Er)/Er$ if $Er > 0$ ; $IE = Ee$ if $Er = 0$		
	Correct correspondences improvement index	[0;n]	$n$
	$CI = (Ce - Cr)/Cr$ if $Cr > 0$ ; $CI = Ce$ if $Cr = 0$		

Ce: number of correct correspondences found in the matching process applied to enriched ontologies.  
 Cr: number of correct correspondences found in the matching process applied to reference ontologies.  
 Ee: number of correspondences expected for a matching process applied to enriched ontologies.  
 Er: number of correspondences expected for a matching process applied to reference ontologies.

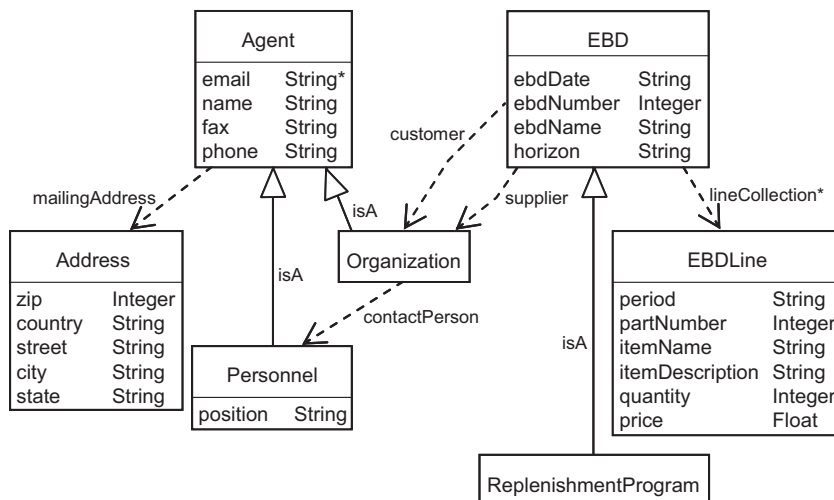
industry called Dairy Packaging that provides packaging to a dairy industry called DairyNY Cooperative. Information the Business Process Management system of each enterprise needs to perform cross-organizational processes is interchanged through Electronic Business Documents. More details about the case study can be found in [25].

**4.1. Hypothesis**

In Peer-to-Peer communication models, enterprises interchange Electronic Business Documents (EBD) to perform cross-organizational processes. EBDs are XML-based standard documents where semantics of information included in them can be defined through ontologies (EBDO). Business processes of each enterprise have to be aligned for achieving semantic interoperability. This requires finding correspondences between the business ontology of each enterprise and the EBDO [25]. These correspondences are obtained by performing an ontology matching process. This process has to deal with heterogeneities caused by both different terminology and knowledge structures, and different contexts in which entities are considered. Then, the quality of the matching process is affected by the quality of the input ontologies.

**4.2. Ontology evaluation**

Consider two ontologies,  $EBDO_{base}$  (Fig. 1) and  $EBDO_{enriched}$  (Fig. 2), which differ in the set of classes, relations, and axioms that use to represent entities stored in an EBD. The  $EBDO_{enriched}$  was obtained from the  $EBDO_{base}$  after applying an enriching process



**Fig. 1.**  $EBDO_{base}$ .

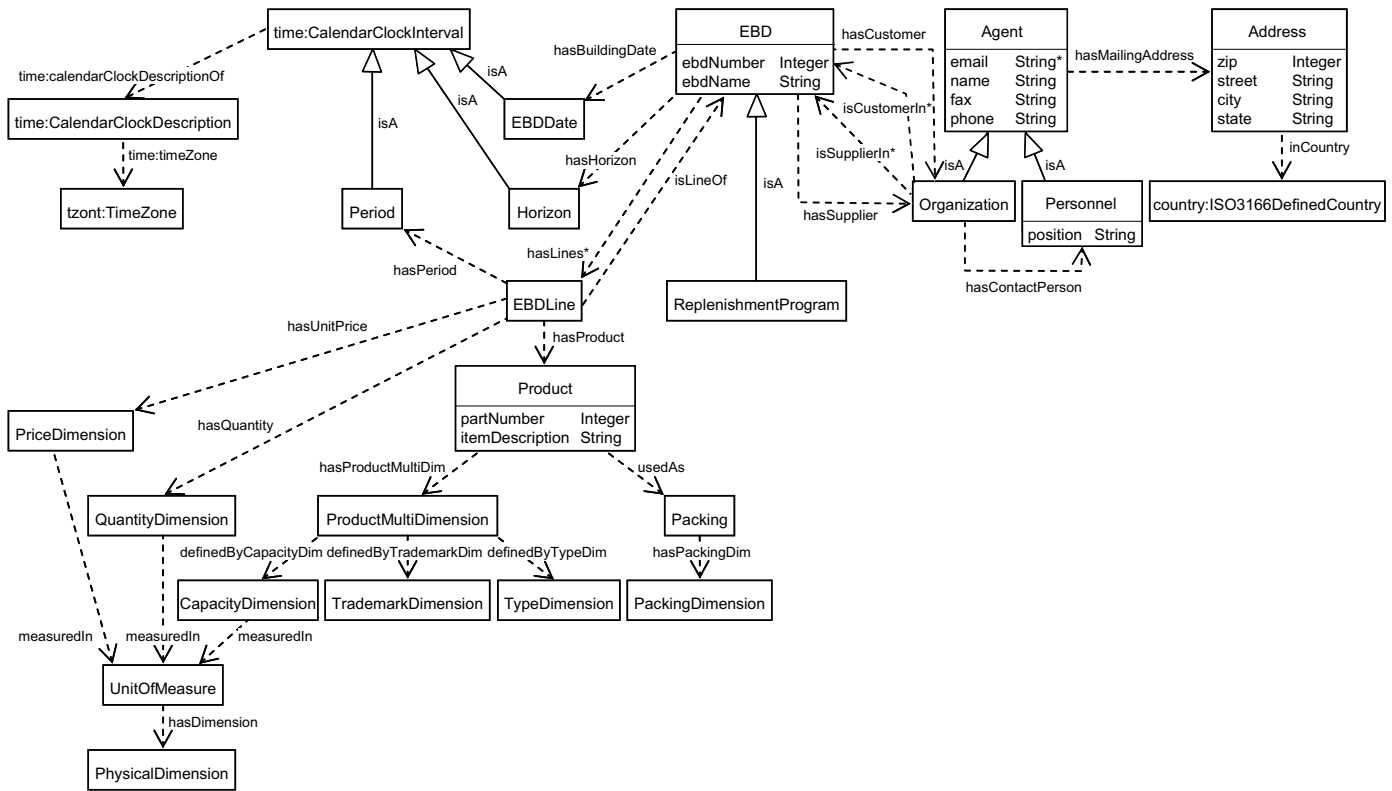


Fig. 2.  $EBDO_{enriched}$ .

Table 9  
Measures for OR1.

Question	Measure	$EBDO_{base}$	$EBDO_{enriched}$
Q1.1	Machine-readable language	1	1
Q1.2	Entities definition capability	1	1
	Relations definition capability	1	1
	Features definition capability	1	1
Q1.3	Potential reasoning	1	1

Table 10  
Measures for OR2; completeness.

Question	Measure	$EBDO_{base}$	$EBDO_{enriched}$
Q2.1.1	NSC	0.00	0.11
	EUR	0.00	1.00
Q2.1.2	DRR	1.00	0.89
Q2.1.3	SPO	1.00	0.58
Q2.1.4	ESPO	0.00	0.33
Q2.1.5	$Coverage(O_c;F_c)$	0.80	1.00
	$Coverage(O_{rc};F_{rc})$	0.30	1.00
	$Coverage(O_{ri};F_{ri})$	0.30	1.00
	$Coverage(O_i;F_i)$	0.80	1.00
Q2.1.6	$Coverage(O_{fc};F_{fc})$	0.67	1.00
Q2.1.7	$Coverage(O_{dfc};F_{dfc})$	0.00	1.00

[25]. Consider also a third ontology  $EBDO_{dairy}$  (Fig. 3) that represents entities of a dairy industry.

AgreementMaker [32] tool was used to calculate the matching between  $EBDO_{base}$  and  $EBDO_{dairy}$ , and between  $EBDO_{enriched}$  and

Table 11  
Measures for OR2; conciseness.

Question	Measure	$EBDO_{base}$	$EBDO_{enriched}$
Q2.2.1	SIC	0.00	0.00
	SII	0.00	0.00
Q2.2.2	RR	0.00	0.00
	ORR	0.00	0.00
	RIR	0.00	0.00
Q2.2.3	$Precision(O_c;F_c)$	1.00	1.00
	$Precision(O_r;F_r)$	1.00	1.00
	$Precision(O_i;F_i)$	0.93	1.00
	$Precision(O_j;F_j)$	1.00	1.00

$EBDO_{dairy}$ . Measure values for  $EBDO_{base}$  and  $EBDO_{enriched}$  according to framework OntoQualitas are summarized in Tables 6–12.

$Alignment_{enriched-dairy}$  refers to the result of the matching process between the  $EBDO_{enriched}$  and the  $EBDO_{dairy}$ . In order to calculate IE and CI, the  $Alignment_{base-dairy}$ , which refers to the result of the matching process between the  $EBDO_{base}$  and the  $EBDO_{dairy}$ , was considered as the reference alignment.

### 4.3. Results analysis

Analyzing results summarized in Tables 9–15, it can be concluded that the quality of the matching process between  $EBDO_{enriched}$  and  $EBDO_{dairy}$  was improved regards to the matching between  $EBDO_{base}$  and  $EBDO_{dairy}$ . This allowed verifying the hypothesis of the empirical study, which states that the quality of the matching process is affected by the quality of the input ontologies. That is, the enriched ontology improves the matching.

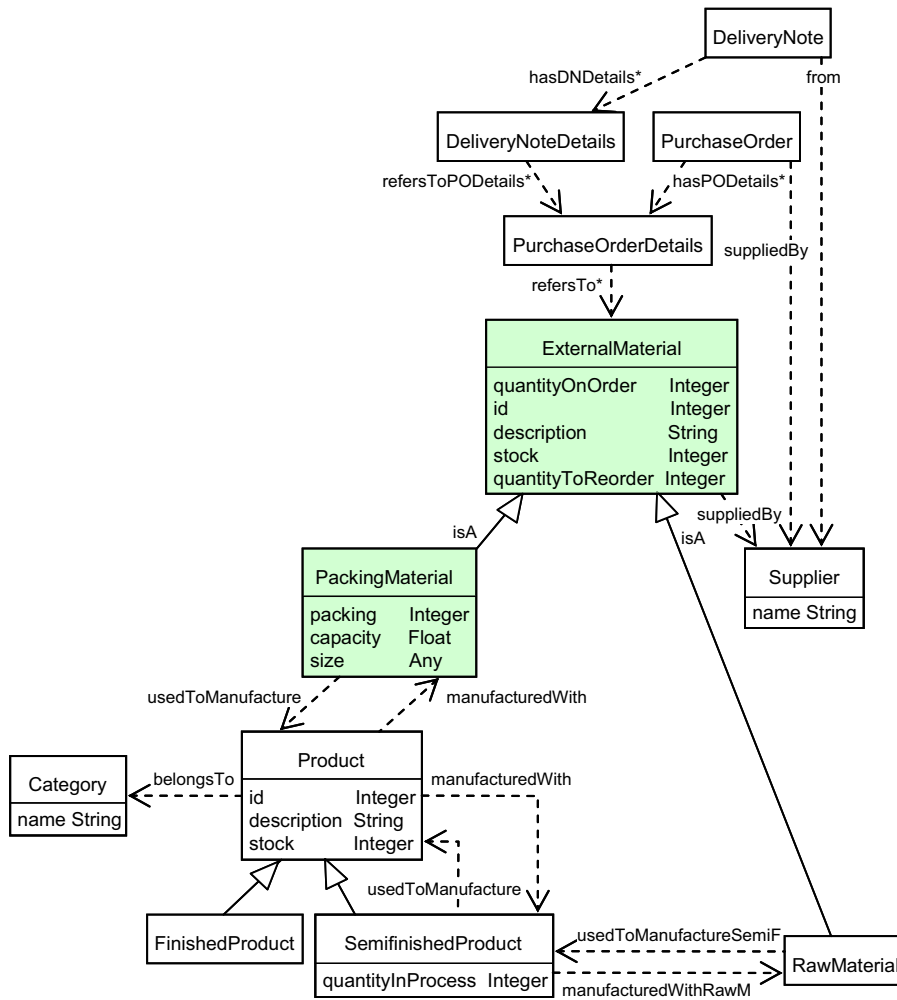


Fig. 3. EBDO<sub>dairy</sub>.

Table 12 Measures for OR3; syntactic correctness.

Question	Measure	EBDO <sub>base</sub>	EBDO <sub>enriched</sub>
Q3.1.1	SL	1.00	1.00
Q3.1.2	R	0.20	0.53

Table 13 Measures for OR3; semantic correctness.

Question	Measure	EBDO <sub>base</sub>	EBDO <sub>enriched</sub>
Q3.2.1	<i>I</i>	0.96	0.93
Q3.2.2	<i>CL</i>	3.52	1.54
Q3.2.3	<i>Cycles(0;0)</i>	0	0
	<i>Cycles(0;1)</i>	0	0
	<i>Cycles(0;n)</i>	0	0
	Subclass partition with common instances	0	0
	Subclass partition with common classes	0	5
	Exhaustive Subclass partition with common instances	0	0
	Exhaustive Subclass partition with common classes	0	0
	Exhaustive Subclass partition with external instances	0	0

Table 14 Measures for OR3.1–OR3.7; representation correctness.

Question	Measure	EBDO <sub>base</sub>	EBDO <sub>enriched</sub>
Q3.2.4	PE	1.00	1.00
Q3.2.5	PU	0.67	1.00
Q3.2.6	PR	0.30	0.60
Q3.2.7	PCS	0.50	1.00
Q3.2.8	PCM	0.42	0.75
Q3.2.9	PCC	0.00	1.00
Q3.2.10	PCc	0.73	0.73

Table 15 Measures for OR3; pragmatics.

Question	Measure	Alignment <sub>enriched-dairy</sub>
Q3.3	IE	2
	CI	2

Regarding results, some peculiarities should be analyzed:

Q2.1.2:  $DRR = 0.89$  due to EBDO<sub>enriched</sub> imports ontologies of country codes ISO 3166 [33] and OWL-Time [34]. OWL-Time has 28 object properties, but only 21 of them have domain and range. This is the reason why the value of measure  $DRR < 1$ .



Q2.1.4:  $ESPO = 0.33$  due to  $EBDO_{base}$  has not disjoint classes and subclasses ( $CDSC = 0$ ), then  $CNCA = 0$  and  $ESPO = 0$ . Ontology of country codes has  $CDSC = 1$  and has not closure axiom. OWL-Time has 2 disjunctions among classes with multiple parents, none of which has the status of closing because they have other children who do not share the same parents.

Q3.2.1:  $I = 0.93$ . However, it must be considered that  $EBDO_{enriched}$  imports ontologies of country codes and OWL-Time, and of the 7 words without meaning in WordNet, only one belongs to  $EBDO_{enriched}$ . Considering only  $EBDO_{enriched}$  elements, regardless imported elements, the interpretation measure of this ontology is  $I = 0.97$ , leading to the conclusion that  $EBDO_{enriched}$  is more interpretable than  $EBDO_{base}$ .

Q3.2.2:  $EBDO_{enriched}$  is clearer than  $EBDO_{base}$  (the closer to 1 is the CL value, the clearer the ontology). An issue of this measure is that compound nouns are unregistered in WordNet, which produces a distortion of its value, since ontologies use compound nouns for naming both terms and properties.

## 5. Conclusions and future work

Based on a top-down approach, framework OntoQualitas proposes evaluation measures for an ontology used for information interchanges. Measures answer questions that help to explore the various aspects of requirements, which contribute to the ontology quality. A requirement is considered to be satisfied when its associated questions are answered and the answers are in a specified range.

OntoQualitas is intended to provide information about the strengths and weaknesses of the ontology, and not to get an overall score. It is therefore not considered the possibility of assigning weights to measures.

In regard to software product quality, SQuARE [35] addresses software product quality in three principal phases of software product life cycle: product under development (internal software quality), product in operation (external software quality), and product in use (quality in use). OntoQualitas covers the first and the last category, since it addresses static attributes of an ontology, such as the size or the compliance with the coding rules, and the capability of the ontology to enable heterogeneous information systems to achieve the information interchange without misunderstandings.

Regarding question Q3.1, the calculation of syntactic correctness, measures of relations richness (RR) and attribute richness (AR) [12] could be useful. However, regarding to the AR it is assumed that the more data properties are defined the more knowledge the ontology conveys [12]. This assumption is relative. There are approaches, such as [25], that reserve the use of data properties to the representation of non-measurable features that cannot be considered entities in themselves. Other features can be represented by means of a set of classes, object properties, and axioms. This fact does not imply that the ontology conveys less knowledge. Optimal values of RR and AR have to be empirically derived based on the information to be interchanged. Exploratory experiments will be conducted to study these optimal values.

Regarding question Q3.2.2, the calculation of the Clarity measure should be improved. An alternative would be to apply a similar strategy to that used by [36] to handle compound nouns. This question will be analyzed in future work.

In Peer-to-Peer communication models between heterogeneous information systems, an autonomous peer can manage simultaneous relations with different peers. In this situation, it is also important to evaluate the ability of an ontology used to interchange information with a given peer as to adapt to changes in order to be used in another relation. To this aim, the ontology should be able to be extended further to describe more fine-grain concepts and relations. This could be defined as requirement OR4: Ontology expandable. Regarding this requirement, the following question could be formulated: Do the set of entities defined allow for future definitions of subclasses? Measures for answer this question will be analyzed in future work.

In this paper, measures for complexity of ontology, such as width and depth, were not considered because they are not relevant for the information interchange, at least as they are actually defined. In the future, the desirability of developing measures to evaluate this aspect of the ontology will be assessed.

The five main design principles proposed for enriching the representation of the interchanged information ( $EBDO_{enriched}$ ) allowed improving the quality of the matching process between dairy industry and  $EBDO_{enriched}$ . In this way, the enrichment process seems to deal with heterogeneities caused by both different terminology and knowledge structures, and different contexts in which entities are considered. Experimental researches will be conducted to study this hypothesis.

## Acknowledgements

The authors are grateful to UTN, CONICET, and ANPCyT for their financial support.

## References

- [1] V. Fortineau, T. Paviot, S. Lamouri, Improving the interoperability of industrial information systems with description logic-based models – the state of the art, *Computers in Industry* 64 (2013) 363–375.
- [2] N. Chungoora, R. Young, G. Gunendran, C. Palmer, Z. Usman, N. Anjum, A. Cutting-Decelle, J. Harding, K. Case, A model-driven ontology approach for manufacturing system interoperability and knowledge sharing, *Computers in Industry* 64 (2013) 392–401.
- [3] J. Jung, Semantic business process integration based on ontology alignment, *Expert Systems with Applications* 36 (2009) 11013–11020.
- [4] S. Khouri, I. Boukhari, L. Bellatreche, E. Sardet, S. Jean, M. Baron, Ontology-based structured web data warehouses for sustainable interoperability: requirement modeling, design methodology and tool, *Computers in Industry* 63 (2012) 799–812.
- [5] G. Wang, T. Wong, X. Wang, An ontology based approach to organize multi-agent assisted supply chain negotiations, *Computers & Industrial Engineering* 65 (2013) 2–15.
- [6] E. Simperl, M. Mochol, T. Bürger, Achieving maturity: the state of practice in ontology engineering in 2009, *International Journal of Computer Science and Applications* 7 (1) (2010) 45–65.
- [7] J. Brank, M. Grobelnik, D. Mladeni, A survey of ontology evaluation techniques, in: *Proceedings of the Conference on Data Mining and Data Warehouses (SiKDD 2005)*, Ljubljana, Slovenia, (2005), pp. 166–169.
- [8] A. Burton-Jones, V.C. Storey, V. Sugumarana, P. Ahluwalia, A semiotic metrics suite for assessing the quality of ontologies, *Data & Knowledge Engineering* 55 (1) (2005) 84–102.
- [9] R. Colomb, Quality of Ontologies in Interoperating Information Systems, Technical Report 18/02 ISIB-CNR, National Research Council, Institute of Biomedical Engineering, Padova, Italy, 2002. Available from: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.100.2321&rep=rep1&type=pdf>.
- [10] B. Stvilia, A model for ontology quality evaluation, *First Monday* 12 (12) (2007).
- [11] K. Supekar, C. Patel, Y. Lee, Characterizing quality of knowledge on Semantic Web, in: V. Barr, Z. Markov (Eds.), *Proceedings of the Seventeenth International Florida Artificial Intelligence Research Society Conference*, AAAI Press, Miami Beach, Florida, USA, 2004.
- [12] S. Tartir, I.B. Arpinar, M. Moore, A.P. Sheth, B. Aleman-Meza, OntoQA: metric-based ontology quality analysis, in: D. Caragea, V. Honavar, I. Muslea, R. Ramakrishnan (Eds.), *Proceedings of IEEE Workshop on Knowledge Acquisition from Distributed, Autonomous, Semantically Heterogeneous Data and Knowledge Sources*, IEEE Computer Society, Houston, TX, 2005, pp. 45–53.
- [13] A. Duque-Ramos, J. Fernández-Breis, R. Stevens, N. Aussenac-Gilles, OQuARE: a SQuARE-based approach for evaluating the quality of ontologies, *Journal of Research and Practice in Information Technology* 43 (2) (2011) 159–176.

- [14] I. Zaihrayev, Towards Peer-to-Peer Information Management Systems, (PhD thesis), International Doctorate School in Information and Communication Technology, University of Trento, Trento, Italy, 2006.
- [15] J. Euzenat, P. Shvaiko, *Ontology Matching*, Springer-Verlag, London, 2007.
- [16] J. Davies, R. Studer, P. Warren, *Semantic Web Technologies: Trends and Research in Ontology-based Systems*, John Wiley, London, 2007.
- [17] A. Gal, P. Shvaiko, Advances in ontology matching, in: T. Dillon, E. Chang, R. Meersman, K. Sycara (Eds.), *Advances in Web Semantics. I: Ontologies, Web Services and Applied Semantic Web*, Springer-Verlag, Berlin, Heidelberg, 2009, pp. 176–198.
- [18] W3C Recommendation, OWL 2 Web Ontology Language, Structural Specification and Functional-style Syntax, 2nd ed., 2012 Available from: <http://www.w3.org/TR/2012/REC-owl2-syntax-20121211/> (accessed on 11.12.12).
- [19] J. Yu, J.A. Thom, A. Tam, Requirements-oriented methodology for evaluating ontologies, *Information Systems* 34 (8) (2009) 766–791.
- [20] N. Guarino, C. Welty, Evaluating ontological decisions with ontoclean, *Communications of ACM* 45 (2) (2002) 61–65.
- [21] A. Lozano-Tello, A. Gómez-Pérez, OntoMetric: a method to choose the appropriate ontology, *Journal of Database Management* 15 (2) (2004) 1–18.
- [22] G. Pierra, Context-explication in conceptual ontologies: the PLIB approach, in: R. Jardim-Gonçalves, J. Cha, A. Steiger-Garção (Eds.), *ISPE CE, A.A. Balkema Publishers*, 2003, pp. 243–253.
- [23] S. Izza, Integration of industrial information systems: from syntactic to semantic integration approaches, *Enterprise Information Systems* 3 (1) (2009) 1–57.
- [24] J. Jung, Service chain-based business alliance formation in service-oriented architecture, *Expert Systems with Applications* 38 (3) (2011) 2206–2211.
- [25] M. Rico, M.L. Caliusco, O. Chiotti, M.R. Galli, An approach to define semantics for BPM systems interoperability, *Enterprise Information Systems* (2013), <http://dx.doi.org/10.1080/17517575.2013.767381>.
- [26] Y. Zhao, K. Sandahl, Potential advantages of semantic web for Internet commerce, in: *Proceedings from ICEIS 2003: The 5th International Conference on Enterprise Information Systems*, Angers, France, (2003), pp. 151–158.
- [27] A. Gómez-Pérez, Evaluation of ontologies, *International Journal of Intelligent Systems* 16 (2001) 391–409.
- [28] M. Horridge, *A Practical Guide to Building OWL Ontologies Using Protégé 4 and CO-ODE tools*, edition 1.2, The University of Manchester, 2009 Available from: [http://owl.cs.manchester.ac.uk/tutorials/protegeowl/tutorial/resources/ProtegeOWLTutorialP4\\_v1\\_2.pdf](http://owl.cs.manchester.ac.uk/tutorials/protegeowl/tutorial/resources/ProtegeOWLTutorialP4_v1_2.pdf).
- [29] N. Guarino, Towards a formal evaluation of ontology quality, *IEEE Intelligent Systems* 19 (4) (2004) 74–81.
- [30] G. Guizzardi, *Ontological Foundations for Structural Conceptual Models*, (PhD thesis), University of Twente, Enschede, The Netherlands, 2005.
- [31] Wordnet, <http://wordnet.princeton.edu/>.
- [32] I.F. Cruz, C. Stroe, M. Palmonari, Interactive user feedback in ontology matching using signature vectors, in: *IEEE 28th International Conference on Data Engineering*, IEEE Computer Society, 2012, 1321–1324.
- [33] DAML, *Ontology*. Available from: <http://www.daml.org/2001/09/countries/iso-3166-ont>.
- [34] J.R. Hobbs, F. Pan, An ontology of time for the semantic web, *ACM Transactions on Asian Language Information Processing (TALIP)* 3 (1) (2004) 66–85.
- [35] ISO/IEC 25000:2005, *Software Engineering – Software product Quality Requirements and Evaluation (SQuaRE) – Guide to SQuaRE*, International Organization for Standardization, Geneva, Switzerland, 2005.
- [36] S. Castano, A. Ferrara, S. Montanelli, G. Racca, Semantic information interoperability in open networked systems, in: *Semantics of a Networked World, Semantics for Grid Databases, First International IFIP Conference, ICSNW 2004, Paris, France, June 17–19, Revised Selected Papers*, Springer-Verlag, Berlin, Heidelberg, 2004, pp. 215–230.



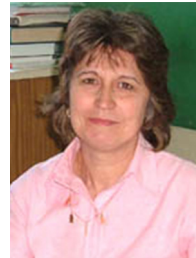
**Mariela Rico** obtained her degree in Information System Engineering in 1995 and her PhD in Information Systems from Universidad Tecnológica Nacional – Facultad Regional Santa Fe (UTN-FRSF), Santa Fe, Argentina, in 2011. From 1995 to 2004, she worked as a consultant and software engineer. From 1995 to 2011, she was an Assistant Professor in Operation Research, and since 2011 she is a full Associate Professor at FRSF. At present, she is working as a research scientist at the I+D Center in Information System Engineering (CIDISI) at UTN-FRSF. Her research activities include decision support system, e-collaboration, context, and ontology with a special focus on semantic interoperability and ontology evaluation.



**María Laura Caliusco** received her degree in Information System Engineering in 1999 and her PhD in Information Systems from UTN-FRSF, Santa Fe, Argentina, in 2005. She is now a full Associate Professor at FRSF. Since 2009, she has been working for National Council of Scientific Research (CONICET) as an assistant researcher. At present, she is working as a research scientist at the I+D Center in Information System Engineering (CIDISI). Her interests are primarily in the areas of e-collaboration, e-government, context, and ontology with a special focus on ontology evaluation, and ontology-aware information systems.



**Omar Chiotti** received his degree in Chemical Engineering in 1984 from UTN – Facultad Regional Villa María, Córdoba, Argentina, and his PhD in Chemical Engineering from Universidad Nacional del Litoral, Santa Fe, Argentina, in 1989. Since 1984, he has been working for CONICET as a researcher. He has been a full Professor of Information System Engineering at UTN-FRSF since 1986. At present, he is the director of the I+D Center in Information System Engineering (CIDISI) and the director of the PhD in Engineering. His current research interests include knowledge management, business process integration, ontology, and multi-agent systems.



**María Rosa Galli** is a professor of information system engineering and industrial engineering at Universidad Tecnológica Nacional (UTN) – Facultad Regional Santa Fe (Argentina). She received her degree in Chemical Engineering in 1983 from Universidad Nacional de Mar del Plata, Argentina, and her PhD in Chemical Engineering from Universidad Nacional del Litoral (Santa Fe), in 1989. Since 1984, she has been working for de Argentina's National Council of Scientific Research (CONICET) as a researcher. Currently, she is member of CIDISI Research Center in Information System Engineering. Her current research interests include knowledge management systems, ontology, e-collaboration, semantic interoperability, and multi-agents systems.