Ontology: Several Theories on the Representation of Knowledge Domains

Maria Luiza de Almeida Campos* and Hagar Espanha Gomes**

*Federal Fluminense University, Information Science, R. Lara Vilela, 126,
São Domingos CEP: 24210-590, Niterói, Rio de Janeiro CEP: 24210-590, Brazil,
<maria.almeida@pq.cnpq.br>
**Conselho Nacional de Desenvolvimento Cientifico e Tecnologico, Brasilia, DF, Brazil,
<hagarespanhagomes@gmail.com>

Maria Luiza de Almeida Campos holds a PhD in information science and is Associate Professor at Fluminense Federal University, Department of Information Science, Niterói, Rio de Janeiro, Brazil. Her areas of interest are knowledge organization, models and theories of knowledge representation, terminology and foundational ontologies.

Hagar Espanha Gomes holds a degree in librarianship and documentation from the National Library Foundation, specializing in master's and bibliographic research by the Brazilian Institute of Bibliography and Documentation, and a MSc in Documentation. Her areas of interest are classification, terminology, information architecture and representation and information retrieval.

Campos, Maria Luiza de Almeida and Hagar Espanha Gomes. 2017. "Ontology: Several Theories on the Representation of Knowledge Domains." *Knowledge Organization* 44(3): 178-186. 28 references

Abstract: Ontologies may be considered knowledge organization systems since the elements interact in a consistent conceptual structure. Theories of the representation of knowledge domains produce models that include definition, representation units, and semantic relationships that are essential for structuring such domain models. A realist viewpoint is proposed to enhance domain ontologies, as definitions provide structure that reveals not only ontological commitment but also relationships between unit representations.

Received: 23 October 2016; Revised: 22 February 2016; Accepted 9 March 2017

Keywords: knowledge entities, knowledge domains, knowledge organization, knowledge representation, concepts, ontology, relationships

1.0 Introduction

Information searching in a quick and precise way has been a perennial problem. On one hand, each view of information seeking comes with a set of relevant concepts and ways of defining and naming them. On the other hand, different modes of representing and relating concepts exist, but in limited models that attempt to express semantics as users do. Information needs and the increasing use of computer resources have changed to a certain degree the role of the information professionals when facing this reality and its many facets. If in the past information professionals working in libraries acted as intermediaries to seek information, today this role is partially performed by computerized systems. Thus, these systems must seek information with the intelligence demanded by users.

More recently, the use of computerized systems supported by thesauri and metadata has required a shift in the role of information professionals who must now interact with such systems to support users' demands and assist in the construction of new information languages to support those demands. Nevertheless, thesauri and controlled vocabularies still have quite restricted semantics from the viewpoint of the machine. The tools control meanings that must be processed only by humans, not by machines. Computers have no intelligence; they are limited to processing instructions, and they are programmed by persons that will act upon metadata. Subject data, if correctly characterized in metadata, can provide unambiguous and precise retrieval.

In this scenario, new tools for information representation with a high degree of formalism have been developed to provide machine processing that can solve user problems as well as other issues regarding information needs. But, to achieve this, new tools and formalisms are not sufficient *per se*. Information professionals must reinvent their roles as intermediaries, intermediaries not between user and information but rather between users' demands and the machine, corresponding to users' perspectives and with concepts and terms familiar to them. In other words, this new intermediary must be able to translate concepts describing the information to formalisms that machines understand. This is, no doubt, a more complex role. One must accommodate the different viewpoints of users and their jargon, and until now formalisms have had limited semantics to handle this, since it must accommodate the complex within the simple with a minimal loss of expressiveness. Inconsistent domain models arise when professionals are not sufficiently aware of knowledge entities that constitute the knowledge domain.

This context needs deeper study on theories and methods regarding the construction of classificatory structures for knowledge and information representation and retrieval. These studies require the development of more consistent strategies and approaches for conceptual modeling of knowledge domains so that high-quality representation is achieved. These models are currently present in semantic tools such as taxonomies and ontologies. Ontologies require a consistent conceptual structure that reveals interaction among their elements as a system of knowledge organization.

A fundamental element in the construction of a conceptual model of a domain is the determination of the representation unit to build the structure of its domain, that is, the nature of the representation elements to be employed in the taxonomy of a given domain. The many theories of representation adopted in the construction of domain models are supported by philosophical undercurrents, not always explicated by its authors, and this leads to consequences when modeling a domain, including the understanding of what may be named as an instance, a class, and even a concept, among other elements intrinsically related to the concept of concept.

Consequently, it is useful to problematize some aspects of the representation unit in a domain representation, such as what knowledge may be ascertained as a concept. From such discussions, semantic relations can also be introduced, since two essential items are necessary for understanding the representation of a domain: node and edge. The role of definitions in domain ontologies must also be emphasized. In this analysis, many of the aspects of domain modeling will be discussed, such as the representation unit, relations among such units, and the ontological commitment stated in their definitions.

2.0 Unit of representation in ontologies

In ontology, several philosophical threads support the understanding of what can be considered a representation unit. Stances such as realism, conceptualism, and nominalism are associated with real-world conceptions that lead the ontologist to adopt a determined perspective directly reflected in the model of the domain being defined.

Universals have concerned philosophers throughout the history of humankind. In the field of knowledge representation and related disciplines, including those involved with development of ontologies, such an issue should be discussed with a goal of consistency between representation models. Advocating for this position, Smith (2004) believes that ontologies developed to support research in natural sciences should adopt not concepts but universals and particulars existing in reality and apprehended by scientific laws. And, depending on how these universals are considered from the modeler's viewpoint, he or she may adopt a realist, a conceptualist or a nominalist position (Smith 2004).

Realism is the philosophical position related to how one recognizes reality in theories of knowledge or metaphysics. In both cases, realism is not opposed to nominalism, but to conceptualism. For realists, there are universals and particulars: the former may be defined as "invariant patterns," also called characteristics or essences, existing with entities of reality. The latter, particulars, are instances of those universals that exist in the real word of space and time. Each universal may be exemplified by a plurality of particulars. So, according to Smith (2004, 79), they are "the instances of such universals which exist in the real world of space and time. The term universal then signifies what the corresponding instances, for instance all whales, all enzymes have in common …. Universals and their instances are joined here by a symbiotic relationship: the (one cannot live without the other) one cannot exist without the other." According to realist viewpoints, relationships of similarity (what instances have in common) would exist even if there were no one to observe them. Relationships of similarity between instances do exist because the same universal is instantiated in each of a given individual form. Universals are organized in trees of genus and species, connected through subtype relations (is a). So, according to a realist approach, concepts may be understood as universals.

Opposing realism, conceptualists assume that no universals exist in reality, only in our minds. Different persons may share the same general concept that, in this sense, exist as multiple examples in different minds. Nevertheless, conceptualists do not generally accept that a concept in one's mind may correspond to a universal or invariant in reality, rather that every concept is related to its instances in only one *ad hoc* way: "things" are constructed by individuals in their minds. More precisely, concepts in general are treated as if all of them are *ad hoc*. Conceptualists' approaches for concepts thus may be understood as general ideas.

Nominalists, according to Mora (2001), assume no universals exist in reality or in our minds, but only general terms exist. These are simple, singular tags of *ad hoc* sets of particulars or events.

Traces of these three philosophical positions are generally found in writings on ontologies and terminology. However, according to Klein and Smith (2010), since they are not clearly distinguished, contradictions have been identified in the relevant literature.

Attempting to clarify the use of the term "concept" as unit of representation in ontology, Smith (2004), and Grenon and Smith (2011) invoke some definitions of concepts found in the literature. Thus, Smith (2004) introduces the principle of universals. This principle is introduced to emphasize the need for evidence, which philosophically supports authors when they refer to concepts in different contexts that are unclear in their texts. Smith's theory of representation suggests a realist approach for ontologies. This approach must provide understanding about what entities do exist in reality—reality here is meant in the most general possible sense including, for instance, not only molecules and planets, but all literary work, laws and historical periods. Objects of ontological investigation belong to a domain and are entities of the first order of that domain, more so than concepts in individual minds (specialists in particular) that study such a domain, or terms used (by specialists, in particular) to refer to a given domain and its elements.

However, concepts and terms may not completely be a matter of psychological, and linguistic, domain ontologies but as entities of first order (as individuals). For instance, in philosophy, Grenon and Smith (2011) and followers argue that many entities to be categorized in ontologies are really concepts. In ontologies of scientific domains, their units are representations of entities as they exist in reality, and these units should be called universals, to which correspond general terms employed in scientific writings, particulars being those instances of universals that exist in the real world of space and time.

In this sense, Smith makes a distinction between ontology based on a realist position and others. The first may be defined as an ontology built with terms that refer exclusively to types that correspond to a given scientific theory. The second one includes artificial representations such as in a taxonomy; the representation units referring to some combinations of types, classes, and certain relationships among them. As a consequence, an ontology based on realism refers to a system of types, while a system of concepts refers to concepts (Klein and Smith 2010). Taking into consideration that the terminology of ontology is not established in a consistent way within the cycle of informatics and terminology, Smith (Smith 2004; Klein and Smith 2010) published fundamental principles to distinguish their theoretical proposal from others. For a better understanding of those principles, according to Smith, systems of concepts are systems of meanings, while systems of representation in ontologies are related to entities of the real world, both those investigated by natural sciences (for example, cells or electrons) and those existing in administrative domains.

His philosophical approach deals with issues that emphasize the roles of definitions, meaning, and reference to build domain ontologies. This approach provides, through careful examination and logical analysis, identification of ambiguities and mistakes when structuring the elements of an ontology, resulting in consistent representation of a domain aiming at logical reasoning. According to Arp and Smith (2008), ontologies should not be built through automatic mechanisms, but by humans, since until now there are no automatic mechanisms with the necessary consistency for natural language processing. Their experience has shown that, given the problems found, the re-examination of the product of natural language processing is a task that rarely achieves soundly structured results. Automatic techniques produce nets of associated terms called lexical networks. These networks are not the most appropriate to represent the structure of a domain. Terminological and lexical information based on links of co-occurrence are useful for some kinds of retrieval but do not offer a reliable representation of a domain.

One limitation of the lexical approach is the lack of interoperability among systems: when the same term belongs to many systems, there is no guarantee that it refers to the same object. An important criterion for quality ontologies is the modularity principle, which should be followed as much as possible for each ontology, as it helps conversion of several ontologies into one; domain ontologies of neighboring fields of knowledge should be built together.

The realist approach, according to Grenon and Smith (2011), is the only stance able to make coherent progress when linking different systems of terminology (for instance, in different languages). Two principles, then, are fundamental: the principle of relevance and the principle of modularity. The principle of relevance means that it is possible to identify entities and their characteristics that belong exclusively to a selected domain; for instance, if the object of the ontology is the philosophy of Bertrand Russell, only entities of this domain will be selected and not of his biography as a political activist; the principle of modularity states that inclusion of new elements in a system in a consistent way must obey each step of subdivision of a class.

Smith (2004) states that concept refers exclusively to a general term, the meaning of which must be only the one agreed upon by individuals of a given disciplinary field of knowledge. The theory of representation, according to arguments presented by Smith is supported by the philosophical current of scientific realism, since the meaning of universals depends on the agreement of qualified individuals in a certain knowledge field.

A useful discussion on concepts as representation units in knowledge domains is Smith on Wüster's (1981) general theory of terminology. According to Smith, Wüster's theory of concepts was destined for failure, and it produces major confusion with heavy consequences in medical informatics projects. Smith argues that concepts are not clearly defined and that definition is seldom viewed as the meaning resulting from a cognitive process, for instance, as a form of knowledge in the mind of a specialist or as term in a system. Problems then arise because simultaneous attempts are made to use the same term for navigating relationships between different entities of the real world (Klein and Smith 2010).

In Wüster's theory, a concept is understood as a mental concept, idea, or thought (Aristotle's *noesis*) that respects a mind state of specific individuals, a state that may be evoked by the use of a correspondent general term. For Smith, this approach is based on a psychological position in which concepts are mental entities, analogously related to ideas and beliefs. This point of view cannot be properly accepted when proposing a standardized terminological system for a domain, since a high degree of arbitrariness and diversity occurs when concepts are constructed in individual minds. Plus, researchers that adopt this position must explain why concepts and their characteristics are either creatures of mind or properties of objects in the world. Thus, it is not clear for individuals creating terminologies whether their elements are representations of ideas in the mind of individuals, meanings of words, knowledge of specialists as a consensus, or types of entities in the world (Klein and Smith 2010; Smith, Ceusters and Temmerman 2005).

Smith does not accept or see as appropriate the use of knowledge units adopted by ISO (2000) to designate concepts. For him, the best formula would be consensual, meaning to avoid psychological connotations. His justification is based on his understanding that it is possible to have an agreed-upon meaning of a term without correspondence to a knowledge unit. For example, a unicorn does not correspond to an entity of reality upon which knowledge may be obtained. Nevertheless, knowledge units may not necessarily have relationships with its existence in reality, providing that a tacit agreement may be accepted in a universe of discourse of a knowledge domain (Dahlberg 1978a).

According to Dahlberg's (1978a) concept theory, knowledge is scientific knowledge when there is agreement about the meaning of a unit within a universe of specialists' discourse. She proposes those terms, as she too disagrees with Wüster's unit of thought, which connotes a psychological view. ISO (2000) adopts knowledge units following arguments in her concept theory. In her paper "Ontical Structures," Dahlberg (1978b) offers a theoretical approach for the construction of a realist concept system, following Vollmer, Campbell and Popper. Vollmer's thesis (1975, quoted in Dahlberg 1978b, 28) is that "every recognition of reality is of a hypothetical nature, and there is a world which is independent of our consciousness, which is structured and which is self-contained. This world is partly recognizable and understandable by perception, thinking and intersubjective science."

For Dahlberg (1978b, 28), that argument is essential to problems of the organization of knowledge because it helps "overcome the existing idealistic approach according to which the formation of knowledge about the world takes place exclusively through the a priori given perceptive forms (*Anschauungsformen*) of space and time and through the thinking and reasoning forms (*Denk*- und *Verstandesformen*)." Like Smith, Dahlberg agrees that Wüster's definition of concept as a unit of thought is inappropriate, since it presupposes a psychological approach with knowledge staying with the individual as a subjective concept. As per Vollmer (1975, quoted in Dahlberg 1978b, 10), there are three distinctive stages of cognition, namely, perception (sub-conscious structure), pre-scientific cognition and scientific cognition. These distinctions contribute to understanding of Dahlberg's conditions set by her statement of concepts as knowledge units.

For Dahlberg, common languages of domains develop first in two stages, namely, perception and pre-scientific cognition. Concepts named and employed with words derived from common language are mental and intellectual property of each human being. In this stage, concepts are understood as units of thought. "Thinking is a cognitive process, depending on the totality of a person's perceptions and experiences, his reflections upon and verification of correct applications of his concepts" (Dahlberg 1978b, 11). But, when speaking of objective knowledge, or that related to scientific domains where scientific knowledge is established, we speak of a knowledge about reality "then it must follow that scientific cognition is concerned with concepts that can be presented in a formal manner" (23). These formal concepts are not those informally acquired by a child; they are scientific concepts, elements or units of cognition, objective/scientific knowledge.

How does one acquire objective knowledge? That is the question proffered by Dahlberg (1978b, 11):

> There is certainly a long way from knowledge about the visible to knowledge about the invisible and immaterial, but it all starts by relating thoughts to the actually existing objects. From statements about his natural surrounding, from experiments and measurements, from counting and inferring man has proceeded to formulate laws of nature and to apply these laws to his fields of activity. Any of his statements relating to a reality experienced or measured creates a 'knowledge element', which may also be regarded as a primitive or a basic concept; And collection of such statements or knowledge elements referring to an object of reality or a verifiable object of thought may be termed a knowledge unit.

Concepts, then, according to Dahlberg (1978a), are established following this sequence: 1) selection of an item of reference of a reality experienced and measured; 2) making true predications about that item, i.e., verifiable statements; and, 3) assignment of a term. A concept may be represented by a triangle with a referent at the top, representing general objects as well as individual ones existing in verifiable scientific reality. Besides referents, the second vortex constitutes characteristics that can be inferred from the referent, from the context in which it is taken, and a third vortex, the denomination, i.e., the verbal expression used for communication. It should be emphasized that what Dahlberg (1978b) considers as_concept characteristics should not be mistaken for referent properties such as: "well defined," "difficult to define," "be too specific," "belonging to other categories," etc. Only at the level of concepts do they become characteristics. There are so many characteristics as there are true statements on the referent in a given context. Among so many possibilities, two kinds of characteristics must be distinguished: the necessary/essential and the accidental. General concepts are defined by the essential characteristics; the specific and individual concepts are described with the accidental ones, in addition to the essential ones (Dahlberg 1978b, 15).

Characteristics of concepts provide elements for conceptual definitions. Such definitions not only are product of the speech or discourse of a group, but fundamentally a conceptual construction of this group, according to agreement among its elements, producing a concept system with categories/facets working as aggregating elements. Dahlberg goes farther with the use of definition; besides being a concept description, it provides not only fixing the concept content but also its position in the respective concept system. It should be noted that even ontologies with realist bases, as supported by Klein and Smith (2010, 436), consider definition as a specification of a concept (i.e., of the agreed meaning of a term), by means of a descriptive statement or a formal expression that serves to differentiate it from other concepts.

Despite the recognition of the importance of definitions, no directions exist on how to use them when structuring domain ontologies. In other words, how do we construct taxonomies that are the basic structures of any domain ontology? As Klein and Smith (2010) have already stated, lexical networks are not appropriate for structuring a domain, but no one orientation is given as to what methodology to follow. As an alternative, Dahlberg (1978b) develops principles of classification for the arrangement of concepts when structuring a domain (and not as a simple list), since concept characteristics support the systematization of concepts. When she proposes a scientific method for concept identification, she meets the need of ontology developers for precise description of referents. In this respect, she takes one step forward toward the construction of definitions, and her model reflects a consensual conception of reality upon a social object, namely, an onomasiological approach, adequate enough for definitions in ontologies and for construction of consistent domain models. According to this approach, the resulting structure reflects how a referent is conceptually built by selecting, highlighting and illuminating some aspects and consequently minimizing or even hiding others. In other words, by illuminating one facet of being, the onomasiological approach produces a version of a referent that constitutes the way a determined group of specialists interact with the referent. This allows a systemic view on the referent where categories have an aggregative role.

3.0 Relationships and definitions in ontologies domain modeling

Relationships deserve special attention from ontology developers. Among the many types, the more familiar ones are hierarchical and partitive, but there may be more in a domain. If for realists some terms need not be defined (primitive terms) in a given context, this does not mean that one does not have resources in natural language to describe concepts represented by terms according to definition patterns that support machine processing as well as their use in data banks. And, several kinds of relationships are made explicit in the description (definition).

Relationships may be classed into two groups: logical relationships and ontical relationships. The first ones are abstract, and they occur among concepts. The second ones occur among objects contiguous in time or space. A more accurate description of the representation elements reveal several other kinds of relationships that are associated with different kinds of entities categorized as endurants (Bittner, Donnelly and Smith 2004, 38):

> Note that individuals, universals, and collections have different temporal properties. Individuals can gain and lose parts. (For example, organisms gain and lose cells.) Universals gain and lose instances. (For example, the universal human being gains or loses instances every time a person is born or dies.) Collections, on the other hand, are identified through their members and thus cannot have different members at different times.

Such details aim at more accurate treatment of data. In some specific domains, however, there is no need for definitions since terms employed have the same meaning among specialists in the field; a concept system requires definitions to assure uniform use of terms in data banks aiming at data sharing and interoperability. Although Smith's examples refer to the biomedical field, they may be generalized. So, a system of terminology has narrow links with ontologies.

Terminology seems to be essential for ontology developers. One of the most operative initiatives in the field is the group Gene Ontology Consortium (2015) that gathers several ontologies, including the popular Gene Ontology (GO). It comprises a controlled and structured vocabulary on the genome domain, and it is a tool for representing and searching information related to genes and products related to every species to be used in genomic annotation. Despite its alphabetic arrangement, GO presents graphs for three genomic aspects, namely, cellular component, cellular function, and biological process. Such graphs complement the limitation of the alphabetic arrangement and help researchers locate concept aspects and attributes, relations among them and find correspondent terms in GO's vocabulary; such a structure still helps find failures in hierarchies. The GO project contains descriptions (definitions) of gene products in the databases of the Gene Ontology Consortium; careful analyses have been made to eliminate inconsistencies. One reason for such inconsistencies seems to be lack of sound orientation for writing definitions. The recommendation offered is a (Gene Ontology Consortium 2015): "textual description of what the term represents and references to information source;" this is not sufficient for an accurate description. The type of definition to be settled should be a logical definition in the format *genus proximum/differentia specifica*.

Another reason for such inconsistencies in GO may be the limitation to only the relationships "type-of" and "part-of;" this is problematic since they are not always consistently used. Besides "is-a" and "part-of" relationships, one finds other relationships such as "has-part," "regulates," and "regulated by," but they are not always sufficient for accurate definitions (Bittner, Donnelly and Smith 2004; Smith and Kumar 2004). Semantic inconsistencies in definitions hamper their use in several circumstances. Yet, it is very strange that in GO there are orientations for standard definitions beginning with verbs in gerund form or in present tense.

Smith and Kumar (2004) analyze definitions in GO, and they find that criteria are needed if a language has to be structured so that information content might be formally expressed and automatically extracted by tools that support logical reasoning. Studies developed on integration of domain ontologies in the field of bioinformatics (Campos et al. 2013; Campos 2007) reveal the existence of several semantic inconsistencies in the OBO Foundry and the Open Biological and Biomedical Ontologies that were created as a collaborative experiment among scientifically based ontologies that also integrate the GO Consortium.

Inconsistencies of a different nature found (Campos 2007; Campos et al. 2013; Smith, Williams and Schulze-Kremer 2003; Ogren et al. 2004; Smith, Ceusters and Temmermanc 2005; Kohler et al. 2006), are related here, namely: 1) hierarchical: failure in the structure of concepts; relevant terms missing in the hierarchy; 2) relational: few relationships available to describe relationships among terms and to express domain knowledge; 3) definitional: inadequate text not explicating characteristics and their relationships, preventing/blocking formalism; lack of patterns for definitions; and, 4) contextual: lack of documentation of ontologies about their scope, objective and subject. Concerning definitions, it is important to have explicit relationships that are richer than those in systematic structure. Methodological procedures are needed, and literature related to ontologies does not mention them, but information science may provide the answer to that, as we will see later. One example is the definition of mitochondria in Gene Ontology Consortium (2015): "A semiautonomous, self-replicating organelle that occurs in varying numbers, shapes, and sizes in the cytoplasm of virtually all eukaryotic cells. It is notably the site of tissue respiration."

According to this definition, if we proceed to a detailed analysis, not only related concepts are presented but also the proper relationship. Three of them are explicit in that ontology: "it_is" a semiautonomous, self-replicating organelle; "occurs_in" in varying numbers, shapes and sizes in the cytoplasm of virtually all eukaryotic cells; "site_of" tissue respiration. This way of analysis may complement or render explicit relationships expressed in the taxonomy of a given subject field.

Study and development of standard relationships then become goals aiming at interoperability. Currently, OBO develops principles for accurate definitions. The site of OBO relationships lists a collection of relations of a general level to be applied by several domains; it intends to provide standards among ontologies of the OBO Foundry and OBO Library. OBO relationships are applicable in several domains (domain-neutral relations) and include (https://code.google.com/p/obo-relations): "part of," "has part," "realized in," "realizes," "occurs," "contain process," "inheres in," "bearer of," "participates in," "has participant," "is role of," "has quality," has role," "derives from," "derives into," "location of," "contained in," "contains," "located in," "boundary of," "member of" and "has member."

Ontobee is another initiative; it is a linked-data server designed for ontologies. It aims at data sharing, visualization, searching, integration and analysis with almost 382 object-property terms used in 150 ontologies (Xiang et al. 2011). Definitions able to support automatic procedures are then a goal to be achieved by ontology.

Returning to Dahlberg, her concept theory proposes an analytical method for concept description. She looks for a scientific method to gain understanding of concepts that leads to accurate and precise description needed for ontology development. She gives an example of a step-by-step construction of predicates leading to a category:

What is a "weekly newspaper?"
a weekly newspaper is a newspaper
a newspaper is a periodically appearing document
a periodical appearing document is a document
a document is a carrier of information
a carrier of information is a carrier
a carrier is a material object

Another example of a general term, not necessarily scientific, is proposed by Dahlberg, with a detailed analysis of a referent (Dahlberg 2000, quoted in Gnoli, Marino and Rosati 2006):

> A bell is a concave container; it is normally made from metal, wood, glass or clay; when it swings it is beaten by an overhanging clapper inside or by an external hammer to produce an intense sound; it is used as a sonorous instrument of resounding solid material to produce vibrations; it has a characteristic form that is dependent of the use in the cultural environment and the material it was made from; its face may be straight, convex, concave, hemispheric, with a cylindrical or tulip form; it has a transverse section that may be circular, square, tetrangular, elyptic or polyhedric; it has a huge geographic distribution; in general it has a very distinguished and defined cultural function; it is used as a convocation (for instance, to religious functions) and with musical purposes.

One can observe that predications include several aspects of observation upon the referent. Depending on the purpose of its use or goal, it is possible to select the necessary aspects. Another example of an accurate and precise term related to mental disease is shown in GO (Ceusters and Smith 2010, 15): "MANIFESTATION OF A MENTAL DISEASE = def. a BODILY FEATURE of an ORGANISM that is (a) deviation from clinical normality that is the *realization of* a MENTAL DISEASE and is (b) observable." Capital letter characters are characteristics that are also concepts, and as such they are elements of description; relationships are in italics. Characteristics are shown in capital letters; it means that they are also defined.

According to Dahlberg (1978a, 1981, 1983), logical definition is the most appropriate to issues discussed here because of its structuring character (see example above). Following the Aristotelian principle of *genus proximum et differentia specifica*, hierarchical relations are made explicit. Dahlberg (1978a) claims that a list of verbs should be worked, and it is clear that what concerns ontology developers also concerns researchers of information science, especially on the development of semantic tools such as taxonomies and faceted terminology systems.

The starting point in an onomasiological approach is an item of reference in a special discourse. In the field of knowledge organization, the approach of Dahlberg's (1978a, 1978b) analytical, referent-oriented concept theory is onomasiological, and it provides a method for definition that must be better investigated, particularly its appropriation in domain ontologies. The referent-oriented principle presupposes that each concept refers to something (a concrete or abstract object) and is analytical, because an analysis of referents identifies their conceptual characteristics thus building concepts analytically and in a structured concept system.

In the social sciences, Riggs (1989a, 1989b, 1996) adjusts Dahlberg's concept theory to disciplinary peculiarities, and he proposes an intermediary model called onomantics that contributes to knowledge organization. According to onomantics, definitions (characteristics that constitute concepts) are compared to identify concepts thus creating terms. Emphasizing referents, he demonstrates that concept theory suits several fields of knowledge.

The onomasiological approach has as significant feature: the use of criteria that allows a consensual conception of reality on a social object thus reflecting the way a referent is conceptually built. This special way of apprehending this conceptual dimension takes into consideration that language is an instrument for the construction of a given reality, and the use of the language approved by a group of experts provides, underpins and modifies social processes. This conception is based in the postulate of the constructivist functionality of language with some of its elements in constructionist lexicology. According to this postulate, language performs, along with other functions (communicative, interactive, etc.), the fundamental role of the construction of referents of speech.

According to Riggs (1989a, 1989b, 1996), the field of linguistics contains a variety of connected disciplines, such as semantics, sociolinguistics, psycholinguistics and language planning that can claim all of them as being associated with the knowledge field of lexicology. On the other hand, related to philosophy one can find logic, philosophy of science and classification research. Within this context, a subfield emerges that focuses on concept analysis, including relations with science in what refers to theoretically significant knowledge units as well as to relations of empirical observations as regarded by operational or denotative criteria. Riggs named this subfield "conceptology," and it is open to different philosophical approaches not necessarily needing to belong to any of them.

The starting point of the onomasiological approach is the item of reference in a special discourse. This does not mean that a semasiological process is adopted, because relationships of meanings in a concept system are not taken into consideration. When taking an item of reference in a given subject context, that item is already the designation of a referent, the object of analysis, resulting in its definition.

4.0 Conclusion

As seen above, explicit theoretical bases for structuring domain ontologies is essential to support consistent decisions when modeling them in a consistent way. Consequently, importance should be given to knowledge of the several theories of representation of knowledge domains. Moreover, knowledge of the principles for definitions in a given ontology is required in order for knowledge discovery by automatic mechanisms to be possible. Methodological strategies to elaborate such statements depend on the assumed theoretical bases.

We believe that there is no right or wrong model, but rather that one serves a given purpose, duly defined and explicated, with suppositions we support. Concept definition is one of the stages of building domain ontologies, as it helps in making consistent decisions. Therefore, we take into account the importance of possessing a wide knowledge of the several theories of representation in those subject fields involved with the representation of knowledge domains. Additionally, we believe that defense of one or other theory is circumstantial, since it depends on the specific or contextual purpose.

References

Arp, Robert and Barry Smith. 2008. "Function, Role, and Disposition in Basic Formal Ontology." *Nature Precedings*. http://precedings.nature.com/documents/1941/version/1/files/npre20081941-1.pdf

Bittner, Thomas, Maureen Donnelly and Barry Smith. 2004. "Individuals, Universals, Collections: On the Foundational Relations of Ontology." In *Formal Ontology in Information Systems: Proceedings of the Third International Conference FOIS November 4-6, 2004 Turin, Italy*, ed. Achille C. Varzi, Laure Vieu. Amsterdam; Washington, DC: IOS Press, 37-48.

Campos, Maria Luiza de Almeida, Maria Luiza Machado Campos, Alberto M. R. Dávila, Hagar Espanha Gomes, Linair Maria Campos and Laura de Lira e Oliveira. 2013. "Information Sciences Methodological Aspects Applied to Ontology Reuse Tools: A Study Based on Genomic Annotations in the Domain of Trypanosomatides." *Knowledge Organization* 40: 50-61.

Campos, Maria Luiza de Almeida. 2007. "Integração de Ontologias: o Domínio da Bioinformática." *RECIIS – R. Eletr. de Com. Inf. Inov. Saúde. Rio de Janeiro* 1: 117-21. http://www.reciis.icict.fiocruz.br/index.php/reciis/article/view/896

Ceusters, Werner and Barry Smith. 2010. "Foundations for a Realist Ontology of Mental Disease." *Journal of Biomedical Semantics* 1, no. 10: 1-23. doi:10.1186/2041-1480-1-10

Dahlberg, Ingetraut. 1978a. "A Referent-oriented Analytical Concept Theory of Interconcept." *International Classification* 5: 142-50.

Dahlberg, Ingetraut. 1978b. *Ontical Structures and Universal Classification*. Bangalore: Sarada Ranganathan Endowment for Library Science.

Dahlberg, Ingetraut. 1981. "Conceptual Definitions for Interconcept." *International Classification* 8: 12-22.

Dahlberg, Ingetraut. 1983. "Terminological Definitions: Characteristics and Demands. *Problèmes de la définition et de la synonymie en terminologie*." Québec: Girsterm, 13-51.

Dahlberg, Ingetraut, 2000. "Strutture Concettuali e Sistematizzazione." In *Categorie e Modelli di Conoscenza*, ed. Giliola Negrini. Milano: Franco Angeli, 171-202.

Gene Ontology Consortium. 2015. http://geneontology.org

Gnoli, Claudio, Vittorio Marino and Luca Rosati. 2006. *Organizzare la Conoscenza*. Milano: HOPS Tecniche nuove.

Grenon, Pierre and Barry Smith. 2011. "Foundations of an Ontology of Philosophy." *Synthese* 182: 185-204.

ISO (International Organization for Standardization). 2000. *Terminology work–Vocabulary. Part 1, Theory and Application. ISO 1087-1.* New York: American National Standards Institute.

Klein, Gunnar O. and Barry Smith. 2010. "Concept Systems and Ontologies: Recommendations for Basic Terminology." *Transactions of the Japanese Society for Artificial Intelligence* 25: 433-41. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3034144/

Kohler, J. et al. 2006. "Quality Control for Terms and Definitions in Ontologies and Taxonomies." *BMC Bioinformatics* 7, no. 212: 1-12.

Mora, J. F. 2001. *Dicionário de Filosofia*. São Paulo: Martins Fontes.

Ogren, P. V. et al. 2004. "The Compositional Structure of Gene Ontology Terms." *Pacific Symposium on Biocomputing* 9: 214-25.

Riggs, Fred W. 1989a. "Information Science and Social Science: The Need for Onomantics." *International Forum for Information and Documentation* 14: 12-21.

Riggs, Fred W. 1989b. "Terminology and Lexicography: their Complementarity." *International Journal of Lexicography* 2: 89-110.

Riggs, Fred W. 1996. "Onomantics and Terminology: Pt. I: Their Contributions to Knowledge Organization." *Knowledge Organization* 23: 25-33.

Smith, Barry and Anand Kumar. 2004. "On Controlled Vocabularies in Bioinformatics: A Case Study in Gene Ontology." *Drug Discovery Today: BIOSILICO* 2, no. 6: 246-51. http://ontology.buffalo.edu/bio/Compositionality_in_GO.pdf

Smith, Barry, Steffen Schulze-Kremer and Jennifer Williams. 2003. "The Ontology of the Gene Ontology." In *AMIA Annual Symposium proceedings, November 8-12, 2003, Washington, DC*. Bethesda, MD: American Medical Informatics Association, 609-13.

Smith, Barry, Werner Ceusters and Rita Temmerman. 2005. "Wüsteria." *Medical Informatics Europe 2005*." Studies in Health Technology and Informatics 116: 647-52.

Smith, Barry. 2004. "Beyond Concepts: Ontology as Reality Representation." In *Formal Ontology in Information Systems: Proceedings of the Third International Conference FOIS November 4-6, 2004 Turin, Italy*, ed. Achille C. Varzi, Laure Vieu. Amsterdam; Washington, DC: IOS Press, 73-84.

Vollmer, G. 1975. *Evolutionare Erkenntnistheorie*. Stuttgart: S. Hirzel Verlag.

Wuester, E. L. 1981. "Étude scientifique générale de la terminologie, zone frontalière entre la linguistique, la logique, l'ontologie, l'informatique et les sciences des choses." In *Textes Choisis de Terminologie: I. Fondéments théoriques de la terminologie*, ed. G. Rondeau; F. Felber. Québec: Girsterm, 57-114.

Xiang, Zuoshuang et al. 2011. "Ontobee: A Linked Data Server and Browser for Ontology Terms." In *Proceedings of the Second International Conference on Biomedical Ontology ICBO 26-30 July 2011 New York, USA*, ed. Olivier Bodenreider, Maryann E. Martone and Alan Ruttenberg. CEUR Workshop Proceedings 833: 279-81. http://ceur-ws.org/Vol-833/paper48.pdf