Looking for ontology in a search engine, one can find so many different approaches that it can be difficult to understand which field of research the subject belongs to and how it can be useful. The term ontology is employed within philosophy, computer science, and information science with different meanings. To take advantage of what ontology theories have to offer, one should understand what they address and where they come from. In information science, except for a few papers, there is no initiative toward clarifying what ontology really is and the connections that it fosters among different research fields. This article provides such a clarification. We begin by revisiting the meaning of the term in its original field, philosophy, to reach its current use in other research fields. We advocate that ontology is a genuine and relevant subject of research in information science. Finally, we conclude by offering our view of the opportunities for interdisciplinary research.

Introduction

Ontology is a subject that has been studied in different research fields and within several domains of knowledge. One can easily find publications on ontology written by authors from philosophy, computer science, and information science with implications for domains such as medicine, biology, engineering, law, and geography. Despite the widespread use of the term across such fields and domains, it is not a trivial task to understand what an ontology is and how it can be useful. An interdisciplinary view may contribute to seeing the whole picture, allowing one to take advantage of the diverse research conducted under the label ontology.

The question of what exists has been studied in philosophy since ancient times, as in Aristotle’s works, although under a different designation. Ontology is considered a branch of metaphysics concerning the study of what categories of entities exist and how each of them is related to each other (Lowe, 2007). The objective is to make sense of the world, of its objects, and of the relations between objects.

In computer science there are a large number of publications on ontology, where the term represents a software engineering artifact (Gruber, 1993). Ontologies are applied in modeling, both in database systems (Wand, Storey, & Weber, 1999) and in knowledge representation systems (Genesereth & Nilsson, 1987). In the context of the former, ontology is seen as a kind of metamodel that one can rely on to promote improvements in conceptual models. In the context of the latter, ontology supports a machine-readable representation of the context aimed at automatic reasoning.

In information science, ontology has been studied since the 19th century, mainly in bibliographic studies for the representation of subjects (Vickery, 1997). This could lead some authors to claim that ontology is a new name for things that have already been studied (Currás, 2003; Gilchrist, 2003; Soerguel, 1999). Evaluating the emergence and intensive use of the term in computer science literature since the 1990s, Vickery concludes that “the problems with which information scientists have for so long been struggling, are now faced by a wider community of knowledge engineers” (p. 285).

A search of Journal of the American Society for Information Science and Technology results in many publications about ontologies, but few of them provide a comprehensive explanation of what ontology is (Fonseca, 2007; Soergel, 1999). What is missing is an approach to the study of ontology and its several facets from an interdisciplinary point of view. In this article, we provide such an approach. We review the meaning of the term ontology in its original field, philosophy, and its use in other applied research fields, such as computer and information sciences, to identify a link between the different uses of the term.

We follow Guarino and Giaretta (1995) when we refer to an ontology—with an indeterminate article and a lowercase initial—as a particular object (an artifact), whereas we use Ontology—with no article and uppercase initial—to refer to the philosophical approach, which deals with the nature and organization of reality. In some cases, the context provides the sense to the reader. We also use the expression
ontological principles to refer to those principles based on the philosophical approach. Examples of these ontological principles are the whole–part theory, types and instantiation, identity, dependence, and unity.

We argue that the label ontology should not be limited to computer science, as it seems to be when one makes a simple search for this term in a search engine (Rittgen, 2012). Rather, we argue that ontological principles are a relevant subject of research in information science. Because studies under the label ontological have been conducted in diverse fields that represent different connotations, we believe that only a comprehensive explanation will allow one to take advantage of opportunities for interdisciplinary research.

The remainder of this article is organized as follows: the second section discusses what Ontology is and the main philosophical approaches, the third section describes how the term ontology is used in other scientific fields, the fourth section discusses additional issues concerning category systems as well misunderstandings about ontologies, and finally, in the fifth section, we conclude by offering our view of the importance of ontological principles for information science and the interplay between different research fields.

What Is Ontology?

The branch of philosophy concerned with the study of reality is known as metaphysics. A fundamental discipline that occupies a prominent position within metaphysics is called Ontology. In the 18th century, German philosophers used the term ontology to refer to a discipline that included notions of being, quality and quantity, and truth and falsehood. In the 20th century, the term was used in philosophical discussions as another name for a scientific logical system (Hennig, 2008).

There is general agreement that Ontology is concerned with what kinds of things can exist. Here, kind means category, a term that was used by Aristotle to discuss the kinds of statements that could be made about an object. A theory of categories is the most important topic of any Ontology. Such theories specify category systems that are structured in hierarchical levels. These systems have in general the form of an inverted tree in which the topmost category is named entity. Anything may be described as an entity of some sort, but the next level of categorization, in general, is a matter of controversy. In addition, there are other difficulties that arise in discussing systems of categories, for example: How many categories exist? Is there a unique topmost category? What is the method by means of which categories should be distinguished?

Before making assumptions according to this or that tradition, one should observe the basic characteristics of the main proposed category systems. Because of their influence in contemporary proposals, at least three such systems are worth examining: the Aristotelian, Kantian, and Husserlian systems (Thomason, 2009). In the remainder of this section, we briefly describe the main tenets of each of these systems.

Other important reasons to study these systems and positions are presented later in the first part of the Discussion section. In addition, essential aspects of the best-known philosophical debate on Ontology, namely realism versus antirealism, are presented to provide a better understanding of the positions presented.

Ontology in Aristotle

Aristotle seems to be the first philosopher to have used the Greek word kategoria as a technical term for predication. Although Aristotle’s writings have different interpretations, followers of his tradition believe that a category system should provide an inventory of the things that exist.

The first category system proposed by Aristotle divided entities in two branches: said-of and present-in. Hence, the following situations are possible: any entity is either said-of another or is not said-of another; any entity is either present-in another or is not present-in another. Entities that are said-of others are called universals, whereas those that are not said-of others are called particulars. Entities that are present-in others are called accidental, whereas those that are not present-in others are nonaccidental. Nonaccidental entities that are universals are described as essential, whereas nonaccidental entities that are particulars are described simply as nonessential. Gathering the possibilities together, one can find a four-fold system of categories and the names provided by Aristotle for each one (Studtmann, 2008):

- Entities that are said-of and present-in → accidental universals
- Entities that are said-of and not present-in → essential universals
- Entities that are not said-of and present-in → accidental particulars
- Entities neither said-of nor present-in anything → primary substances

A second category system was proposed by Aristotle that contains a list of the highest kinds, which are known as categories. Noticing that ordinary objects of human experience fall into classes of increasing generality, one can see indications that the existence of the highest kind is probable. However, Aristotle did not believe in a unique high level, but in 10 of them as shown in Figure 1.

The most important category, substance, can be understood by listing its significant characteristics (Smith, 1997). They can exist on their own; they remain numerically one and the same while having different properties at different times; they can stand in causal relations; they have no proper parts that are themselves substances; they have a continuous existence; and they have no temporal parts.

One might ask what the different sorts of questions that would have been employed to obtain Aristotle’s list of categories are. By “sorts of questions,” we mean different questions that may be asked about something. For example, the question “What is it?” can only be asked of a substance, and only answers describing substances are appropriate (Ackrill,
and plants by the differentia of self-movement. For example, the genus of living things is divided into animals that have a particular differentia and those that do not. For Aristotle, each entity has a fundamental feature called its real essence. Another question, which was posed by Kant, concerns the unsystematic way that Aristotle’s categories might have been chosen, which makes it impossible for one to know whether Aristotle’s list of categories is complete (Jansen, 2008).

### Ontology in Kant

One of the main concepts in Kant is the distinction between analytic and synthetic propositions in the scope of subject–predicate judgments. According to Kant, judgments are complex conscious cognitions that refer to objects either directly through intuitions or indirectly through concepts. Analytic propositions contain the predicate concept a priori in the subject concept, whereas synthetic propositions do not (Hanna, 2006). An example of an analytic proposition is “all triangles have three sides”; an example of a synthetic proposition is “all creatures with hearts have kidneys.”

For Kant, the terms a priori and a posteriori are linked to the recognition that human knowledge cannot be supported only by sensory observation. Accordingly, we have a priori knowledge, which is expressed, for example, in the knowledge of logical truths and mathematical principles. Examples of a priori prepositions are “all bachelors are unmarried” and “3 + 2 = 5” (Kitcher, 2006).

Skepticism about the possibility of discerning the different categories of reality led Kant to try an approach in which the aim was not to catalog things in the world, but to determine which categories belonged to the human category system (Thomason, 2009). For Kant, one can have a priori knowledge by means of categories only if the categories are imposed by the mind on the objects that it knows (Rohlf, 2010).

To arrive at a list of categories, Kant departed from the concept of a judgment in the traditional propositional form of the Aristotelian logical system: that is, subject–copula–predicate, for example, or “S is P.” There are four aspects that are proposed for classifying judgments: quantity, quality, relation, and modality (Wood, 2004). For the aspect of quantity, a subject term may be one of three kinds:

<table>
<thead>
<tr>
<th>Universal</th>
<th>Particular</th>
<th>Singular</th>
</tr>
</thead>
<tbody>
<tr>
<td>All S</td>
<td>Some S</td>
<td>This S</td>
</tr>
<tr>
<td>All dogs have hair</td>
<td>Some dogs have hair</td>
<td>Fido has hair</td>
</tr>
</tbody>
</table>

A similar rationale is applied to the categories of quality, relation and modality.

For quality, the idea is similar to quantity but considers predicates instead. The difference between the negative and the infinitive judgment is that the former negates an affirmative predicate, and the latter negates the affirmative predicate by leaving open other possibilities in the domain. Hence, it would be correct to predicate the number “five” as the negative “five is not green,” but incorrect to predicate it as an infinite, which would mean that “five” is not blue because it has another color (which is truly false).
The book is red  The book is not red  The book is not red, but it is yellow, blue, etc.

For modality, three kinds of copula are proposed to connect subject and predicate: problematical, assertory, and apodictic:

<table>
<thead>
<tr>
<th>Problematical</th>
<th>Assertory</th>
<th>Apodictic</th>
</tr>
</thead>
<tbody>
<tr>
<td>S is possibly P</td>
<td>S is effectively P</td>
<td>S is necessarily P</td>
</tr>
<tr>
<td>The table is possibly brown</td>
<td>The table is effectively brown</td>
<td>The table is necessarily brown</td>
</tr>
</tbody>
</table>

Finally, for relation, judgments can be combined in syllogistic inferences, as is possible in traditional logics. This results in three possibilities:

<table>
<thead>
<tr>
<th>Categorical</th>
<th>Hypothetical</th>
<th>Disjunctive</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Si sP</td>
<td>If S is P, then S is R</td>
<td>S is P or R</td>
</tr>
<tr>
<td>All pets are animals</td>
<td>If a pet is a dog, then it is barking</td>
<td>A dog is barking or a cat is meowing</td>
</tr>
</tbody>
</table>

These aspects are presented in Kant’s table of judgments, which is depicted in Figure 2.

Kant also proposed that, for each entry in the table of judgments, there would be an equivalent entry in a table of categories. Proceeding in this direction, Kant arrived at the 12 correlated concepts of human understanding depicted in Figure 3.

Although it deals with the categories of human understanding, Kant’s theory retains Ontological importance insofar as the categories apply a priori to all objects of cognition. By delineating the concepts that are necessary for cognition, one can acquire knowledge of the categories governing any object of cognition, but not of the things themselves.

**Ontology in Husserl**

The phenomenology developed by Husserl is the study of “phenomena,” meaning the ways in which things appear to us in different forms of conscious experience (Smith, 2005). Husserl teaches us how to perform a “phenomenological reduction” by abstaining from perceiving the natural world and by assuming a transcendental position to describe only pure consciousness. This method would allow one to realize an object as it is experienced, with no concern as to whether it exists. Such consciousness is intentional, in the sense that it is directed at an object (Smith & Smith, 1995).

Husserl makes a distinction between, on one hand, categories of objects, which represent the highest possible kinds of objects, and on the other hand, categories of meaning, by means of which one can think about kinds of objects. The term category has a particular sense in Husserl’s work; it is an entity from the realm of formal essences (see Figure 4). Beginning with an entity called object-in-general, one can reach the highest formal essences through a content removal process called formalization. In contrast, through a process of generalization, one can add content to the nature or essence of an entity and reach the highest material species (Beyer, 2011). There are material entities, such as dog, a dog, and dogs in general, which serve as the basis for the process of formalization toward the purely formal level. Accordingly, upon reaching the formal level, we have formal entities such as something, this something, and something in general.

The process of formalization reflects an organization of reality that encompasses the common characteristics of both objects and kinds of objects, involving part–whole and dependence relationships. It is worth noticing that the term object here does not have the trivial meaning of a physical entity. Acts of consciousness also have part–whole and dependence relationships. In Husserl’s parlance, “formal ontology” is the discipline that studies these types of relations. This discipline allows one, once laws governing certain entities have been determined, to apply the results to other entities of the same nature.

When expanding his theories, Husserl constantly improved his philosophical system by integrating views about Ontology, epistemology, and logic (Smith, 2003). Here, one can notice how Husserl’s studies can contribute to the study of ontology. For example, the aforementioned

---

**FIG. 2.** Table of judgments. Adapted from Kant (1999).

**FIG. 3.** Kant’s table of categories. Adapted from Kant (1999).
notion of intentionality raises several Ontological questions about mental acts, their contents, and the objects toward which they are directed (Siewert, 2006).

In Husserl’s ontology, the major distinctions were made among the realm of fact (I), the realm of essence (II), and the realm of meaning or sense (III). The two first realms, fact and essence, are called domains of knowledge: the realm of fact is the domain of empirical sciences such as physics, biology, and psychology, whereas the realm of essence is the domain of eidetic sciences such as mathematics, logic, and formal and material ontology. The third realm, meaning, constitutes the domain of phenomenology. Below, we briefly explain Husserl’s ontology, focusing on the entities represented in Figure 4 (Smith & Smith, 1995).

The realm of fact (I) includes concrete entities, namely, concrete empirical individuals, concrete empirical states of affairs, and concrete empirical events. In this context, “concrete” refers to real entities that exist in time and space, and “empirical” refers to entities studied by empirical sciences, in contrast to those studied in phenomenology as advocated by Husserl. Individuals (I-1) can be either independent individuals (I-1.1), also called substrates, or dependent individuals (I-1.2), also called moments. Examples of substrates are trees and tables; examples of moments are instances of colors in trees. A moment is actually a dependent part, one that cannot exist unless the whole of which it is a part exists. A state of affairs (I-2) is categorically formed from individuals. Examples of such a configuration would be a table and its brownness, a tree and its greenness, and so forth. In the latter example, the greenness would be both an instance of green and a moment of that tree. Events (I-3) are entities that have duration. An event happens to an individual in a certain instant of time. The most important events are experiences, intentional acts that manifest the consciousness of something.

The realm of essence (II), in opposition to the realm of fact, contains only ideal entities. Ideal entities are nonreal; they do not exist in space or time. Within the realm of essence there are both formal essence (II-1) and material essence (II-2), in addition to region (II-3). Formal essence (II-1) contains entities that are pure forms representing material species. These pure forms are categories. Among the highest categories, those that define the forms of all objects in the world are individual (II-1.1); essence, which encompasses species, quality, and relation (II-1.2); and state of affairs (II-1.3). Formal essences are also the subject matter of formal ontologies. Material essence (II-2) consists of hierarchies of entities of the natural sciences, organized in a genus–species form. Material essence is the subject matter of material ontologies. Regions (II-3) are the highest species in the hierarchy of material species. They are distinguished in nature (II-3.1), which includes all natural things, such as animals and plants; consciousness (II-3.2), which includes all conscious experiences; and spirit (II-3.3), which includes perception, judgment, and imagination.

The realm of meaning or sense (III) contains meanings, that is, how people think about objects and kinds of objects. Meanings are not located in time and space. Meanings are ideal; they are not essences, but the contents of intentional experiences, the ways in which objects are presented in consciousness. Each intentional experience has a content, the central part of which is called sense (III-1).

In addition to the definition of categories as ideal entities in the realm of formal essence, three types of categories are distinguished (Smith & Smith, 1995): (a) linguistic categories, such as names, predicates, and statements; (b) meaning categories, such as propositions and concepts; and (c) Ontological categories, such as species, qualities, and states of affairs. In addition, Husserl conceived a formal ontology discipline analogous to the previously developed “formal

FIG. 4. Husserl’s basic ontology (capitalized letters denote essences or species per se). Adapted from Smith and Smith (1995).
logics” (Johansson, 2004). Formal disciplines are set apart from material ones insofar as the former apply to all domains of objects and are independent of the peculiarities of a specific field of knowledge.

**Realism and Antirealism**

Aristotle used language as a clue to describe Ontological categories. Kant used concepts as a way to approach categories of objects of possible cognition. The goal of the Aristotelian and Kantian category systems was to describe the categorical structure that the world would have according to human thought and language. Husserl provided categories that are descriptive of the highest essences of possible things. Departing from Aristotle’s categories, Husserl does not attempt to provide an inventory of what things actually exist as a matter of fact.

Despite the differences among the influential approaches presented so far, philosophical theories in general have been classified according to two main positions: realism and antirealism. The debate involving these two positions has been ongoing for centuries, and there is no consensus (MacLeod & Rubenstein, 2005).

The term realism has several interpretations in philosophy (Niiniluoto, 1999); but, in general, realism is used to designate the notion that there is a mind-independent physical world. Mind-independent entities are called universals, which are instantiated by objects called particulars. Universals are employed by realists to explain relations of qualitative identity and resemblance among particulars. According to realists, for example, the sentence, “The moon is spherical” is true independent of anyone’s beliefs and linguistic practices.

The first realist proposals came from Plato and Aristotle. For Plato, true knowledge would be permanent and unchanging; because of this it cannot come from ordinary objects, which are subject to constant change. Thus, for Plato, permanent things, the so-called universals, come from a realm of forms apart from the everyday world. In this line of thought, a bed and a table are ideas (or forms) that people had before they were born. Aristotle could not relate Plato’s realm of forms to the particulars of everyday life. Therefore, he proposed that universals should not be separated from reality, but should be common elements present in particulars of the same category. For example, the universal table consists of all characteristics common to all tables.

Nonrealism can take many forms depending on whether the dimension of realism that is rejected is existence or independence of entities. In general, antirealists are classified as nominalists and conceptualists (MacLeod & Rubenstein, 2005).

Nominalists believe that only individuals exist and the problems of identity and resemblance can be solved by thinking about individuals and the relationships between them. They believe that empirical similarities between things are not good criteria for establishing membership in a category or for characterizing a universal. The separate existence of universals, a dualist view exhibited both in Plato’s and Aristotle’s views, is the main point of discordance with realism. Nominalists advocate that universals are unnecessary and that all knowledge comes from particular entities of people’s experience.

Conceptualists deny that individuals suffice to solve the problem, but they do not appeal to universals. Instead, conceptualists explain identity and resemblance by referring to concepts or ideas. In this view, each word of language has a general concept associated with it. The conceptualist view seems to bring back a sort of dualism, similar to the realistic view, that establishes a difference between things and the abstractions of things.

Although these positions are not always clearly stated, one can see that, whereas Aristotle was guided by ontological realism, Kant and Husserl subscribed to some form of antirealism. Kant, in general, is considered an idealist, someone asserting that reality is fundamentally mental, mentally constructed, or otherwise immaterial. Husserl is said to adhere to versions of both realism and idealism (Beyer, 2011).

**Ontology Beyond Philosophy**

Beyond its usage in philosophy, the term ontology has been employed in other fields of research to refer to different things. Even within computer science, a field that uses the term intensively, there is more than one sense for ontology. This section presents a brief description of the main approaches to dealing with an ontology and Ontology, both in computer science and in information science.

**Ontology in Computer Science**

One can easily find publications about ontologies coming from computer science and dealing with domains such as medicine, biology, geography, and law, to mention but a few (Gene Ontology Consortium, 2003; Johansson & LYNÖE, 2008; Koepsell, 2000). In knowledge representation (KR), a subfield of artificial intelligence (AI), the term has been used since the 1960s to mean a general structure of concepts represented by a logical vocabulary. In the 1990s, the term continued to be used within the set of technologies encompassed under the label semantic web, which brought a kind of renaissance to AI by promising to bring the possibility of automatic inferences to the web.

Even considering only the realm of computer science, the study of ontologies is characterized by the coexistence of diverse definitions (Guizzardi, 2005). To deal with the issue of term usage plurality, Smith (2004) distinguishes two types of ontologies. The first relates to how the universe is organized and corresponds to the philosophical approach, and the second concerns the task of conceptualizing a domain. This second fulfills the need to express an ontology as an artifact within the scope of software engineering and KR.
By considering the main activities and agents involved in knowledge representation, one is better able to understand the role of ontology in KR. Software engineers aim to develop declarative systems that contain assertions representing facts governed by rules. An example of a fact is, "New York is a city in the United States," and an example of a rule is, "All people who live in New York live in the United States." This combination of facts and rules composes the so-called knowledge base of a system. A knowledge base is built and maintained by a knowledge engineer, who has the main task of formalizing the knowledge of a group of experts and representing it in a format that allows automatic inferences. In addition, a knowledge engineer also has the duty to improve the knowledge base by acquiring new facts and updating rules as new knowledge is obtained.

As part of this work, one both evaluates facts and identifies their structure, as is done when building a theory. To do this, one has to make generalizations and abstractions, which requires metaphysical insights (Ontology). Therefore, following this line of thought, an ontology is a theory representing the main facts and rules governing a certain part of the reality. The term ontology, thus, is another term to name this theory.

There are other KR approaches that follow a similar rationale but employ more expressive languages of representation, such as modal logics, to build models of a theory (Guarino, 1998). By following an ontological commitment—a partial semantic account of the intended conceptualization of a logical theory (Guarino & Giaretta, 1995)—a knowledge representation language is employed to generate a set of models representing a certain conceptualization of reality. An ontology makes explicit axioms that permit such models to be constrained so that they match, as closely as possible, the models that contain the intended meaning (Figure 5).

A conceptualization C would be represented by a KR language L following an ontological commitment K = <C, I>. This can be better understood by saying that L commits to a domain D through K, from which C is the underlying conceptualization. The variable I represents the interpretation function, which maps elements of D to symbols of a vocabulary V. All models of L that are compatible with K are intended models of L according to K. All variables depicted in Figure 5 are defined according to a traditional AI book (Genereseth & Nilsson, 1987). We do not present here the complete mathematical formula for the precise and formal definition of all entities.

In such a framework, the role of the ontology is to constrain the models of L so that they become intended models I_L. The best ontology corresponds to the closest approximation between models and intended models, as depicted at Figure 6.

One can notice two main meanings for the term ontology in computer science. The first has to do with the use of ontological principles to understand reality and model it. Ontology is used in line with its original role in philosophy, namely, to provide an account of what exists and characterize entities in modeling activities. The second concerns the representation of a domain in a KR language. Ontologies here consist of a set of statements written in a KR language, which can be processed by automatic reasoners.

An example of the first approach is the work on evaluation of models according to ontological principles (Wand & Weber, 1990; Wand, Storey, & Weber, 1999). An example of this second approach is the use of the web-oriented KR language called OWL (Web Ontology Language) to build ontologies within the context of the semantic web (Staab & Studer, 2004). However, the mere use of the term does not oblige an engineer dealing with a computational language to make good Ontological decisions to represent the world. In general, the decisions are taken considering system performance limitations, resulting, in many cases, in a poor representation of reality (Smith & Welty, 2001).

A more detailed view of KR theories goes beyond the goals of this article, but anyone interested can find additional information (e.g., Sowa, 1999, 2000).
Ontology in Information Science

Software engineering projects involving ontologies exhibit many parallels with theories traditionally studied in information science, such as faceted classification, controlled vocabularies, and lexicography. Issues faced by computer scientists in the digital era when building machine-readable vocabularies are issues that information scientists have been advancing for several years (Vickery, 1997).

There are differences and similarities between ontologies in computer science and two other terms widely employed in information science, namely taxonomy and thesaurus. The possibility of constraining natural language seems to be the point of contact among these three sorts of structures (Gilchrist, 2003). In addition, although the term ontology sometimes appears as a synonym for controlled vocabularies, those structures have different objectives (Currás, 2003). Whereas controlled vocabularies aim to retrieve documents, ontologies aim to model a domain to produce automatic inferences. But there are other uses for the term ontology in information science.

Fox (1983) uses philosophical grounds to explain the hypothesis that information has an ontological status similar to the propositions underlying a text; both are atemporal and nonspatial, in addition to being abstract objects. Propositions do not coincide with the sentences of a text, insofar as several sentences can express the same proposition. For example, the sentences, “Mars has two moons,” “Two moons circle Mars,” and “Mars a deux lunes” convey the same proposition, that is, that Mars has two moons.

In this line of thought, the information contained in a document cannot be identified with the text of the document, but with its propositional content. To reach the propositional content, one would conduct a propositional analysis of the information following the principle that “the information carried by sentence S is a proposition appropriately associated with S” (Fox, 1983, p. 84).

This propositional approach originates from the study of logic and of philosophy of language, namely, from the distinction proposed by Frege. Frege used as examples the expressions Morning Star and Evening Star, which have different meanings, but refer to the same entity, the planet Venus. Frege attributed the entities of his theory to the three vertices of a triangle, calling them symbol, sense, and reference. This same distinction was introduced in linguistic semantics by Ogden and Richards (1989) through the triangle of meaning: “a diagram where the three factors involved, every time we declare that something is understood, are put on the vertices of the triangle, with the existing relations among them being represented by the sides” (p. 32).

Dahlberg (1978) presents the Triangle of Concept, which is different from the Ogden and Richards triangle, because the concept relates to the vertices of the triangle, and its meaning might assume many possibilities. Dahlberg (1978) presents the concept as being the sum of the true and essential statements about a referent, and the term (lexicon) as being the communicable and representable form of the concept. In this representation, we can see that the meaning of a term—its verbal form—is derived from the interpretation of a set of true statements (characteristics), which may be attributed to a referent, that is, an object, phenomenon, process, or entity.

In the context of information retrieval systems, Blair (2006) made clear a commitment to the later philosophy of Wittgenstein. Wittgenstein’s approach to categorization is based on the notion of family resemblance, which has two basic assumptions: the members of a taxonomic group share a cluster of similar traits, and those traits do not need to occur in all members of a group and only that group (as is the case with Aristotle). Wittgenstein called family resemblances the relations among entities in a taxonomic group; they are serially related resemblances. For example, there are games in a set that share traits A and B and C; games in a second set share traits A and B and D and F; and games in a third set share the traits D and F and G. The games in the first and second sets directly resemble each other; the games in the first and third sets only indirectly, serially, resemble one another (Ereshefsky, 2001).

The significance of the Wittgensteinean view for the design and use of information retrieval systems lies in the fact that “indeterminacy will be an essential characteristic of many content retrieval efforts” (Blair, 2006, p. 339). Therefore, one should assume the existence of several sorts of indeterminacy and plan strategies to mitigate each case (e.g., semantic ambiguity, category overload, and content representation).

There are several other approaches in information science that, directly or indirectly, address ontological issues. It seems that information science authors do not always explicitly mention the term ontology, although ontological principles appear clearly in the literature in that field. Anyone interested should consult, for example, the work of Ranganathan (1967), Lancaster (1986), Foskett (1985), and Buckland (1991), to mention but a few.

Discussion

We have presented different uses for the term ontology in different research fields. Our survey is not intended to be exhaustive, insofar as it seems impossible to encompass all the possibilities of this subject. Rather, we have briefly surveyed the more influential philosophical category systems and their use in applied fields, with the aim of clarifying their main characteristics. At this point, some issues are worth discussing: the validity of the category system presented so far and the possibility of having a position in the realism–nominalism debate. In addition, some usual confusions about ontologies are pointed out, as a way to foster the understanding of the subject.

Although initiatives for comparing philosophical systems of categories exist (Moss, 1964; Perreault, 1994), such a task is not trivial. As one can notice when looking over the most influential systems, even the term category itself has
Different meanings in distinct systems. When considering the explanatory power of each system, one can question the validity of using only one system. The categorical systems vary to such an extent that it is hard to believe that one complete and exhaustive system can be found (Westerhoff, 2005). Despite this, these systems continue to be employed, and new ones have been proposed in the past 50 years, the majority of them based on the category systems surveyed in this paper. Examples of these new systems are Johansson (2004), Chisholm (1996), Grossman (1983), Lowe (2007), and Armstrong (1989).

New category systems also have been developed to meet specific goals already mentioned, such as those in computer science. In this context, the most currently used systems are the Descriptive Ontology for Linguistic and Cognitive Engineering (Gangemi et al., 2002) and the Basic Foundational Ontology (Grenon, Smith, & Goldberg, 2007). Whereas the former is a high-level (generic) computational ontology inspired by Kant, the latter is based mainly on Husserl and Armstrong. In information science, category systems also have been created with the aim of representing document contents for retrieval. A well-known example in this field is that of Ranganathan (1967).

It is evident that systems used in computer science and information science have different goals. Computer science uses ontology to categorize the world, but emphasizes the reasoning process. The emphasis on reasoning falls within the realm of logics, with the aim of finding out how a pattern of syllogism combines two premises to arrive at a conclusion. In addition, these systems are extremely reductionist, because they have to deal with the limitations of representation and performance of computational systems. In contrast, the main concern of information science is dealing with documents that describe the entities of the world within different domains and from different points of view, not with logical reasoning.

It is worthwhile, at this point, to reinforce the important aforementioned distinction between Ontology in philosophy and ontologies in computer science and information science. The first corresponds to a philosophical discipline, namely that branch of philosophy dealing with the nature and organization of reality; the second and third refer to a particular determinate object, an artifact, which can be an informal conceptual system, a formal semantic account, a specification of a conceptualization, a representation of a conceptual system via a logical theory, a vocabulary used by a logical theory, or a meta-level specification of a logical theory (Guarino & Giaretta, 1995).

Considering the distinctions presented so far, Table 1 summarizes the main interpretations of ontologies in different contexts. As the term theory is used in Table 1, a theory does not always refer to the whole reality, but that it can be used to represent knowledge in a specific domain (Reynolds, 2007).

The last line of Table 1 classifies a catalog, a glossary, and a thesaurus as ontologies according to the view of Smith and Welty (2001), which considers these structures as informal conceptual systems. However, this view is arguable; one can claim that ontology is not a classification system or a controlled vocabulary in the sense that controlled vocabularies are used in information science (i.e., as an authorized or normalized set of topical index terms). Instead, an ontology is a kind of controlled vocabulary in that it specifies the set of predicates that can be used to make statements (i.e., representations) about a resource.

Another issue worth being discussed concerns how one could choose a position within the realism versus antirealism debate. Different positions form a kind of spectrum between the two extremes, and several of them can be considered correct. A definitive answer to this issue would be as complex as the original debate itself, which has been ongoing in philosophy for thousands of years. However, the importance of this issue lies in the fact that, once the characteristics of the main positions are known, one can understand the perspective from which the world is seen.

There are arguments in favor of a philosophical realism point of view in information science. Even subscribing to the very same basic tenet of realists—the metaphysical view that some things exist independent of human thought—Hjørland (2004) does not refer to ontological realism, but to epistemological realism. Although within the scope of the former, the issue is what entities are real and whether they

<table>
<thead>
<tr>
<th>Distinction</th>
<th>Field</th>
<th>What is it</th>
<th>Purpose</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ontology as a discipline</td>
<td>Philosophy</td>
<td>Ontology as a system of categories</td>
<td>Understand reality, things that exist, and their characteristics</td>
<td>Systems of Aristotle, Kant, Husserl</td>
</tr>
<tr>
<td>Ontology as an artifact</td>
<td>Computer science</td>
<td>Ontology as a theory (logic-based)</td>
<td>Understand a domain and reduce it to models</td>
<td>BFO, DOLCE (generic)</td>
</tr>
<tr>
<td></td>
<td>Information science</td>
<td>Ontology as a software artifact</td>
<td>Create vocabularies for representation in systems and generate inferences</td>
<td>OWL (KR language)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ontology as an informal conceptual system</td>
<td>Create controlled vocabularies for information retrieval from documents</td>
<td>Ranganathan’s Classification System A catalog, a glossary, a thesaurus</td>
</tr>
</tbody>
</table>

Note: BFO indicates Basic Foundational Ontology; DOLCE, Descriptive Ontology for Linguistic and Cognitive Engineering; KR, knowledge representation; OWL, Web Ontology Language.
are mind-independent; the latter addresses whether knowledge about the world is possible (Niiniluoto, 1999). It is clear that epistemological realism deals with human knowledge, not the world itself.

Within the scope of information retrieval systems, this central realist position corresponds to the consideration that a given document can be relevant to a certain purpose, independent of a user’s thoughts: “whether or not a certain substance is relevant as a cure for cancer is ultimately decided in medical research, not by asking patients or users of medical services” (Hjørland, 2004, p. 497). This position has consequences for the design and use of retrieval systems, insofar as identifying what is relevant in terms of content corresponds to an engagement in scientific arguments.

On the other hand, there are positions aligned with some sort of antirealism. Wittgenstein, for example, assumes a unique position rejecting both strict realism and nominalism. This position is said to be unique because Wittgenstein’s earlier writings exhibited a position close to realist views, whereas his later writings admitted misgivings about his earlier position (Blair, 2006). Following this line of thought, not all nominalist notions were rejected, but there is an objection to the fact that a nominalist does not consider the importance of context in the meaning of language.

In relation to realism, Wittgenstein’s most important rejection is the claim that one can find a finite number of properties of a thing that allows one to categorize that thing. The most well-known example in Wittgenstein is his attempt to define “games,” which reveals more distinct characteristics than common ones (Wittgenstein, 2009). What then are the essential properties that exist in all games?

Finally, special attention should be given to the possibilities for misinterpretation of the interdisciplinary literature of ontology, which this article intends to minimize. There is a general belief that connects the word formal to the realm of logic. However, in some contexts, the word formal means a “theory of forms,” which is employed to understand the characteristics of objects (Husserl, 2001). Another common conflation occurs when one mixes ontology with epistemology. Whereas the former deals with the characterization of things, encompassing theories for whole–part, identity, instantiation, and so forth, the latter is a theory of knowledge that studies how one can know about the world (Hessen, 1999; Niiniluoto, 1999). Other possibilities for misinterpretation are the result of a lack of distinction among kinds of semantics, for example, when one makes a direct comparison between the semantics of computational ontologies, which is formal semantics, and the semantics of natural language (Almeida, Souza, & Fonseca, 2011).

Summary and Conclusions

In this article, we have revisited different views of ontology with an emphasis on its original field, namely philosophy. We described three influential systems within Ontology: realist, cognitive, and descriptive, which are connected with the theories developed by Aristotle, Kant, and Husserl. Then, we presented an overview of ontologies outside of philosophy, as the term is currently employed in other research fields. We explained the interplay between these fields and discussed some issues related to ontologies. We would now like to offer our conclusions and what we believe are opportunities for interdisciplinary research on ontologies.

Ontology in computer science is used to refer both to a vocabulary expressed in a KR language and a kind of theory where one explains phenomena using facts and rules. The first use corresponds to software, a computational artifact that has to be developed according to implementation guidelines. The second use retains the notion of Ontology from philosophy, of an inventory of things in the world and relations among them in a particular domain based on principles of the Ontology as a discipline. It is in this last sense that ontology has the power to connect several fields.

In information science, ontological principles may be used to support the building of categorical structures for representation of the content of documents. Ontological work in information science is not confined to representation of the subject content of resources, but encompasses representation of the resource as a whole, generally from the perspective of a particular community of users. Ontology in this sense is a genuine and fruitful subject of research in information science, because it holds the potential to explain not only issues regarding the content of a document, but the entire social environment involved in the analysis conducted by information scientists. A detailed explanation of this topic is beyond the goals of this paper, but anyone interested in such a possibility should look into the discipline of social ontology and may consult Zaibert and Smith (2010), Lawson (2008), Tsohatzidis (2007), Tuomela (2007), Bratman (1999), Searle (1995), Gilbert (1989), and Mulligan (1987).

Evidence of the use of ontological principles in information science is presented in this article. Important and difficult issues—for example, what information is—are dealt with from an Ontological perspective. However, one can claim that the majority of references come from other fields and have nothing to do with information science. Even considering only the computational ontologies that are being created in large scale in domains like medicine (Smith et al., 2007), one should see ontologies (ontologies as artifacts) as repositories of scientific knowledge, which are, from another perspective, an object of interest for information science.

In addition, one can consider that the creation of computational ontologies involves modeling, which is “[. . .] the activity of formally describing some aspects of the physical and social world around us for purposes of understanding and communication” (Mylopoulos, 1992, p. 3). Describing aspects of the physical and social world for communication purposes is an activity that falls within the scope of information science (Saracevic, 1999). Again, from another perspective, ontologies as artifacts could also be studied as part of a communication phenomenon.
What connects different fields of research are the ontological principles, which are the common element present in philosophy, computer science, and information science. To reach an interdisciplinary level, one should consider the task of dealing with ontological principles in two stages; the first is a priori and devoted to establishing what kinds of things could exist and coexist in the world, and the second involves the effort to establish what kind of thing does exist according to empirical evidence. The first takes place in philosophy, the second in applied sciences such as information science.

The benefit for philosophy is to have ontological principles tested in practical enterprises, providing a new use for theoretical notions in important fields such as medicine and law. The benefit for applied sciences is to have solid and robust principles, some of them under discussion for more than 2,000 years, to ground their empirical investigations. Within this scenario lies the opportunity for interdisciplinary research in ontologies. Within this context, the literature in information science about ontologies should be built.

One can argue that the rigor of Ontological theories is not always required in some activities conducted in empirical research and may even be an obstacle. We believe that a good example of how to overcome this impression is the relation between natural language semantics and formal semantics. Developments made in formal semantics, which work in a controlled scenario, can be then tried in natural language, in a real-world scenario. Developments in formal ontologies (based on a logical language) can produce results that can later be applied in controlled vocabularies, which deal with natural language.

Finally, one may believe that the existence of different category systems burdens us with the responsibility of choosing which one to use. As we see, these systems can have a realist, cognitive, or descriptive slant, that is, they can be systems that do not share the same principles. However, from the distinction between formal and material ontology, Husserl taught us that we can use different views on the formal level (philosophical) as a basis for understanding material ontologies in applied sciences. This frees one to choose one or another category system using a realist universal or a nominalist state of affairs to explain the material level. Empirical results can determine which would be the better approach.

References


