Ontology Revisited:
Concepts, Languages, and the World(s)

Jari PALOMÄKI1
Tampere University of Technology/Pori
Department of Information Technology

Abstract. In this paper the role of ontology is considered in the context of information systems and conceptual modelling. Firstly we describe information systems and conceptual modelling 
without using the word “ontology”. Then two senses of ontology are presented: philosophical and knowledge representation views. In this connection both Gruber’s and Guarino’s definitions of “ontology” are scrutinized. It is proposed that the kind of terminological shifts concerning the word “ontology”, and practiced e.g. by Gruber and Guarino, have caused more confusion than clarification in the field of information systems and conceptual modelling. Instead of giving a new or a better definition of our own, our view is that the word “ontology” should be used in its traditional philosophical sense only, whereas our aim in the field of information systems and conceptual modelling is to reach conceptual clarity.

Keywords. ontology, conceptual modelling, information systems, knowledge representation

Introduction

In the field of information system the word “ontology” has nowadays at least three different meanings. Firstly it may mean a conceptual model, or a conceptualized knowledge representation. Secondly it may mean a vocabulary or a dictionary containing the basic terms used in conceptualization, i.e., a language. Finally, and in its original philosophical meaning, ontology is a part of metaphysics concerning the universe of discourse and the ultimate building blocks of the world.

A definition of ontology, which has gained some agreement among the information system community, is the following one: “An ontology is a formal explicit specification of a shared conceptualization for a domain of interest”, [1]. In this definition the word “ontology” means a formal language, although the two other possible meanings for ontology are presented as well. This definition has been criticized by Guarino and Giaretta in [2], and elaborated by Guarino in [3], in which he gives the definition of his own: “An ontology is a logical theory accounting for the intended meaning of a formal vocabulary, i.e. its ontological commitment to a particular conceptualization of the world”. We shall scrutinize both of these definitions.

This paper is composed as follows. Firstly we define the basic modelling situation consisting of objects to be modelled, a modeller, i.e. a person who is modelling, a model, and the relationships between these. We continue this process by describing an information system as an instance of a basic modelling situation. Having described information systems we consider a conceptual modelling process, the goal of which is to develop a conceptual schema of the domain of an application. In this context also the nature of concepts and concept theory are considered. The above account of information systems and conceptual modelling are described without using the word “ontology”. Then two senses of ontology are presented: philosophical and knowledge representation views. In this connection both Gruber’s and Guarino’s definitions of “ontology” are scrutinized. Then some reasons are given why we should avoid using the word “ontology” in the area of information systems and conceptual modelling.

1 Corresponding Author: Jari Palomäki, Tampere University of Technology/Pori, Department of Information Technology, Pohjoisranta, P.O.Box 300, FI-28101 Pori, Finland; E-mail: jari.palomaki@tut.fi
1. Basic Modelling Situation

Modelling starts when for some special purposes something is to be modelled. This creates a basic modelling situation, which consists of 1) objects to be modelled, 2) a modeller, who is doing the modelling, 3) a model, which is to be a result of the modelling, and 4) different relationships between these. The basic modelling situation is then a three-place relation: $M(\text{object}(x), \text{modeller}(y), \text{model}(z))$, which is to be read: “an object $x$ is modelled by $y$ as $z$.” From this definition it follows that modelling is always an intentional activity.

The objects to be modelled form the object domain, i.e. the universe of discourse. The objects in the object domain do not have necessarily to be concrete things in space and time of which we are to have immediate sense perception, but can as well be abstract objects consisting e.g. expert’s knowledge. Moreover, the objects to be modelled do not have to exist before the model of it is created. This situation happens, for example, when designing or planning something.

A modeller is the subject of a modelling situation. In most cases a modeller consists of the group of persons having different expertises. The most important task for a modeller is to consider those features of the objects to be modelled which are relevant for the purpose of modelling.

A model is a result of an abstraction that is used to represent the objects in the object domain. Abstraction is an epistemological process, where some relevant aspect of the objects in the object domain is separated. There are also other epistemological processes connected with creating a model, for example, classification, i.e. grouping the objects into classes on the basis of some of common properties; generalization, i.e. arriving at some general notion from the instances; axiomatization, i.e. by giving the basic propositions (truths) from which we can deduce other propositions (truths) the result of which is called an axiom system; etc.

Finally, modelling relations consist of logical and epistemological relationships between the modeller, model, and an object domain. Depending on the order of modelling relations, we can talk about the descriptive and the “normative” use of the model. The model is descriptive, when we start from the object domain, and via the modeller we will end to the model, whereas the model is normative, when we start from the modeller, and via the model we will end to the object domain.

2. Information System

Having defined the basic modelling situation, we shall proceed to an information system. Firstly, any system may be characterized as a set of interrelated components that interact with each others to form a whole, and work together toward a common goal. The four major components of a system are (1) inputs, (2) processes, (3) outputs, and (4) feedback control. An input is anything that enters the system. A system may have one or many inputs. An output is anything leaving the system. Outputs are the goal of the system, the purpose for which the system exists. The outputs of one system can become the inputs of another system. A process performs the work: it transforms an input into an output. The system exchanges inputs and outputs with its environment. Changes in the environment affect the system. To adopt itself to changing conditions the system have a feedback. A feedback reintroduces a portion of the output of a system as an input into the same system, and thus, a feedback is actually a special kind of input. It is used as a control mechanism within the system to indicate the difference between the goals and actual performance of a system, see e.g. [4], [5].

In information system inputs are of two kinds: firstly, they are result of gathering and capturing data to be stored in databases and, secondly, they consist of queries concerning the data stored in databases. Information processing involves converting or transforming queries into valuable information, and it can be done manually or with the assistance of a computer. Processing performed by a computer is actually performed by the computer’s Central Processing Unit (CPU) which, in doing that, interplays with databases and a conceptual schema, sometimes also called the logical schema in relational data models. Resulting information, as an answer to a query, may be used as a feedback to make changes to input or processing activities.

An information system is not isolated from its environment, and so it has also a larger sense. An information system is a set of people, data, and processes that work together to achieve the
common goal of information management. A computer-based information system includes also hardware, which consists of everything in the physical layer, and software, which consists of systems programs that control the hardware and application programs used to produce the required information.

3. Conceptual Modelling and Conceptual Schema

A conceptual modelling consists of analysing and synthesising concepts which are “drawn” from a given object domain, and constructing a conceptual model of it by means of those concepts, [6], [7], [8]. The result of this process is presented in a form of conceptual schema, which provides a conceptual description of the data under inspection. Accordingly, databases are used to store data representing a certain section of the reality, and the classification and organization of that data is described by means of a conceptual schema. A conceptual schema of the object domain, in turn, is used as an input to a support system, e.g. CASE tool, which generates a database of it. The database implementation includes update and query functionalities as well.

In the construction of the conceptual schema at least the following four principles would serve as guidelines, [9], [10], [11]:

1. The conceptualization principle, according to which only conceptual aspects should be taken into account when constructing the conceptual schema.
2. 100% -principle, according to which all the relevant aspects of the domain should be described in the conceptual schema.
3. The formalization principle, according to which conceptual schemata should be formalisable in order to be implementable.
4. The semiotic principle according to which conceptual schemata should be easily interpretable and understandable.

From the principles 1. and 2. it follows that all the relevant aspects of the domain appearing in the conceptual schema are to be conceptual. The principle 3. is essentially a restrictive principle, and it is partly opposite to the principle 4. The motivation for the principle 4. is that the language for representing conceptual schemata should be common for everyone taking part in the modelling process, and be appropriate for fulfilling the task of modelling. A reason why the principle 4. may sometimes be opposite to the principle 3. is that not everyone is familiar with the formalisms used in this area.

In order to express the result of conceptual modelling as a conceptual schema, we need a modelling language. The three most popular notations for information modelling are Object-Role Modelling (ORM), Entity-Relationship (ER) diagrams, and Unified Modelling Language (UML), see e.g. [12]. However, many accounts of conceptual modelling emphasize an intensional aspect concept as opposed to real world things in the domain of objects, which belong to the extensions of those concepts. This means that intensionality should be visible in the modelling language as well, [7], [8].

Let us note that there are two different inquiries each having a right to the name “conceptual modelling”. The first one is when we are actually doing conceptual modelling, and the second one is when we are describing the process of conceptual modelling. The first one is more basic, since in the second one conceptual modelling is conceptually modelled, and hence, it has different status than the first one. In any given discussion, it is easy to fall into confusion through failure to determine to which of the two inquiries the discussion is intended to belong. In this paper we are interested in the second one from which the first one follows.

5. Concepts and Concept Theory

A conceptual model is composed of particular concepts and particular relations between them. A concept theory, in turn, is studying concepts and the relations and operations between them in general. There are numerous theories of concepts, for example, concepts can be thought of as any
of the following: *supersensible entities* as universals, meanings, abstract objects, definitions, or predicates and relations; *mental entities* or *states* as composite images, ideas, thoughts, conceptions, or innate ideas; *neutral entities* joining e.g. words, thoughts, and things; human or animal *skills* or *abilities*, etc., see e.g. [13]. In order to condense this great number of theories without sacrificing too many of their indispensable properties, the following more abstract classifications can be made:

1. *The entity theories of concepts*, according to which concepts are entities such as ideas, images, thought, abstract objects, etc.
2. *The dispositional theories of concepts*, according to which concepts are understood to be the abilities or skills to do something, especially the ability to use language correctly.

We may go even further. One of the major doctrines of both the entity and the dispositional theories of concepts is that a concept, in order to be a concept, must be characterized by a definite set of necessary and sufficient criteria. This is called a *closed concept*, and the requirement that all concepts are to be closed is called the *closure principle of concepts*, [14]. Thus we have abstracted the particular conceptions of concepts to the point where there are just concepts, and whatever form these take, they must satisfy the closure principle of concepts. The advantage of this approach is that developing a concept theory it does not confine us to any particular theory of concepts with its specific accompanying restrictions and special characteristics. Relations between concepts enable us to establish conceptual structures. The most essential relation between concepts is the intensional containment relation, which is based on the intensions of concepts, see [7], [8], [14], [15]. Another important relation between concepts is the conceptual containment relation between concepts, which is based on the conceptual constituents of concepts, and is thus a different relation than the intensional containment, [16], [17].

The idea of the concept of a concept and the relations between concepts serves as the basis for all conceptual modelling, and all conceptual structures are ultimately reducible to concept theory, which is to be understood in the similar way as “all mathematics is ultimately reducible to set theory”, [18].

6. **Ontology: Philosophical and Knowledge Representational Views**

The above account of forming an information systems and doing conceptual modelling is written *without* using the word “ontology”. Nowadays there are two sense of the definition of that word: the traditional one used in philosophy and the more modern one used in the area of information systems, [19]. We will consider firstly the philosophical view and then turn to the area of information systems, where the view is called a knowledge representational view.

6.1. **Ontology: A Philosophical View**

In philosophy ontology is a part of metaphysics, which aims to answer at least the following three questions:

1. What is there?
2. What is it, that there is?
3. How is that, that there is?

The first is (1) is perhaps the most difficult one, as it asks what elements the world is made up of, or rather, what are the building blocks from which the world is composed. A traditional answer to this question is that the world consists of things and properties (and relations). An alternative answer can be found in Wittgenstein’s *Tractatus* 1.1: “The world is the totality of facts, not of things” [20], that is to say, the world consists of facts. Still another answer can be a process ontology proposed by Alfred North Whitehead in [21] and presented e.g. in [22].
The second sense of the word "ontology" is used in research on information systems, artificial intelligence and knowledge representation. One of the best known definitions is Tom Gruber's: system community, see Preface in [23] and [24].

It has gained some agreement among the information discourse. The domain is represented in a language, the set of objects so represented form the universe of machine interpretable meaning of concepts and relations between them. When knowledge of a domain is represented in a language, the set of objects so represented form the universe of discourse.

The second question (2) concerns the basic stuff from which the world is made. The world could be made out of one kind of stuff only, for example, water, as Thales suggested, or the world may be made out of two or more different kinds of stuff, for example, mind and matter.

The third question (3) concerns the mode of existence. Answers to this question could be the following ones, according to which something exists in the sense that:

a) it has some kind of concrete space-time existence,
b) it has some kind of abstract (mental) existence,
c) it has some kind of transcendental existence, in the sense that it extends beyond the space-time existence.

In conceptual modelling the most crucial ontological question is: “What modes of existence may concepts have?” This question is related to the third question (3) above. The traditional answers to it are that

(i) concepts are merely predicate expressions of some language, i.e. they exist concretely, (nominalism);
(ii) concepts exist in the sense that we have the socio-biological cognitive capacity to identify, classify, and characterize or perceive relationships between things in various ways, i.e. they exist abstractly, (conceptualism);
(iii) concepts exist independently of both language and human cognition, i.e. transcendentally, (realism).

If the concepts exist only concretely as linguistic terms, then there are only extensional relationships between them. If the concepts exist abstractly as a cognitive capacity, then conceptualization is a private activity done by human mind. If the concepts exist transcendentally independently of both language and human cognition, then we have a problem of knowledge acquisition of them. Thus, the ontological question of the mode of existence of concepts is such a deep philosophical issue, and part of a reason to adopt the closure principle of concepts is to be more neutral on these questions.

However, if we take an ontological commitment to a certain view of the mode of the existence of concepts, consequently we are making other ontological commitments as well. For example, realism on concepts is usually connected with realism of world as well. In conceptualism we are more or less creating our world(s) by conceptualization, and in nominalism there are neither intensionality nor abstract (not to say transcendental) entities like numbers.

6.2. Ontology: A Knowledge Representative View

The second sense of the word “ontology” is used in research on information systems, artificial intelligence and knowledge representation. One of the best known definitions is Tom Gruber’s:

“An ontology is a formal explicit specification of a shared conceptualization for a domain of interest” [1].

In this definition the word “ontology” means a formal language, or more exactly, by using the language of the Section 4, a conceptual schema, cf. [6]. Although Gruber’s definition has been criticized by Guarino and Giaretta in [2], it has gained some agreement among the information system community, see Preface in [23] and [24].

According to Gruber [1], the reason to use the word “ontology” to refer to a language is that for knowledge-based systems only those items “exist” which can be represented in a language which comes with a formal semantics. Using this formal approach to ontology provides the machine interpretable meaning of concepts and relations between them. When knowledge of a domain is represented in a language, the set of objects so represented form the universe of discourse.
This language dominated view to the world has a strong resemblance with Rudolf Carnap’s empirical approach to semantics and (philosophical) ontology, cf. [25]. He distinguishes between internal and external linguistic frameworks. Meaningful questions are empirical questions about the world, which can be asked only within the linguistic framework, i.e. they are internal questions, whereas the external questions concerning the existence or reality of the systems of entities as a whole are pseudo-questions. That is, the above mentioned ontological questions presented in philosophy are more or less meaningless.

From this, in turn, it follows that if our knowledge of the universe of discourse can only be thought in connection with a language and a conceptual schema, and not as independent of them, then the truth of the world can only be defined internally – and hence, there does not need to be one, but many worlds. Moreover, as far as a “conceptualization” is a private mental process done by a human being, and which aim is to form a concept of something, there cannot be a “shared conceptualization” either. Hence, it seems that from the philosophical point of view Gruber’s definition of “ontology” in [1] as well as in [25] is simply incoherent.

Nicola Guarino, on the other hand, in [3] recognizes the two readings of “ontology”: AI reading and the philosophical reading. He says firstly that he will adopt the AI reading, i.e. the knowledge representation view, and will use the word “conceptualization” to refer to the philosophical one. However, instead of the philosophical view of ontology represented in the Section 5 above, Guarino refers by the philosophical readings to “an ontology as a particular system of categories accounting for a certain vision of the world”.

Now, Guarino’s definition of “ontology” in the context of information systems is the following:

“An ontology is a logical theory accounting for the intended meaning of a formal vocabulary, i.e. its ontological commitment to a particular conceptualization of the world” [3].

In the first reading of this definition we may think that by using the language of the Section 4, the “intended meaning” refers to a conceptual model, in which a conceptual model is a product of conceptualization of the world, and the conceptual model is a mental “vision of the world” with a certain formal structure, (a formal ontology). However, when Guarino starts to formalize his account of “ontology” by using the possible world semantics, we will learn that the “ontology” for him is a set of axioms (a language) \(O\) such that its models approximate as well as possible the intended models of a language \(L\) according to an ontological commitment to a particular conceptualization \(C\) of the world. He also emphasizes that “it is important to stress that an ontology is language-dependent, while a conceptualization is language-independent” [3]. Here, however, the word “conceptualization” means “a set of conceptual relations defined on a domain space”, whereas by “the ontological commitments” Guarino means the relation between the language and the conceptualization.

Thus, in [3] Guarino is using the word “conceptualization” to mean two different things: on one hand the philosophical readings of “ontology”, and, on the other hand, a formal structure \(C = <D,W,R>\), where \(D\) is a domain, \(W\) is a set of possible worlds, and \(R\) is a set of conceptual relations on the domain space \(<D,W>\). A conceptual relation, in turn, is a total (extensional) function from the set of possible worlds \(W\) to a power set of domain \(D\). Still, the standard meaning of the word “conceptualization” is “to form a concept” of something.

To illustrate the above definition of the word “ontology”, (i.e. making it more confusing), Guarino gives a picture, in which this time the word “ontology” refers not to a language \(O\), but to an intended model of the language \(L\), see Figure 1.

---

2 Although Guarino emphasizes intensional aspect of modelling it is to be noted that possible world semantics reduces to an extensional view by using possibly partial, but indeed extensional set-theoretic functions from the possible worlds to extensions of those worlds.
Figure 1. The intended models of a logical language reflect its commitment to a conceptualization, [3].

No doubt the Figure 1 is misleading. It does not describe the situation as Guarino actually writes about it in [3]. So, instead we will propose the following illustration, see Figure 2.

Figure 2. Ontology $O_L$ as a language connected with the language $L$ and its model $M(O_L)$ approaching and including the intended model $IM(L)$ of the language $L$.

In Figure 2 “ontology” $O_L$, is a language associated with the language $L$, and its model $M(O_L)$ includes the intended model $IM(L)$, and which is also included into the all of possible models of $L$, i.e. $M(L)$. Provided that the Figure 2 is more correct illustration of Guarino’s idea, we can make the following observations concerning ontology $O_L$. Firstly, $O_L$ is a language, which is only indirectly connected with a “conceptualization” $C$ as well as an “ontological commitment” $K$, which connects the language $L$ with the “conceptualization” $C$. Secondly, the interpretation function of $O_L$ is missing and no formal structure, i.e. a “conceptualization”, of it is given. Thirdly, in order the model $M(O_L)$ of $O_L$ to approximate and include the intended model $IM(L)$ of $L$, $O_L$ should be a kind of a meta-language of $L$ restricting the interpretation of the language $L$ to the intended model of it, i.e. $IM(L)$. Fourthly, both languages $O_L$ and $L$ are making (in a philosophical sense) ontological commitments to possible worlds, and to the basic domain $D$ consisting of things, properties, and relations.
6.3. So, Why Should We Worry?

There is a wide variety of possible interpretations of the word “ontology”, see [2]:

1. A philosophical discipline
2. A conceptual system
3. A representation of a conceptual system
4. A formal semantic
5. A specification of a “conceptualization”
6. A specification of a logical theory
7. The vocabulary used by a logical theory

The interpretation 1. is a traditional philosophical use of the word “ontology”, whereas all the other interpretations 2.-7. belong to the knowledge representation view and appear in the literature of information systems. Now, for example, an “ontology” for Guarino as well as for Gruber was seen to be a language-dependent, which takes in account of its explication both concepts and the world(s). Since it is possible to describe the process of information systems and conceptual modelling *without* using the word “ontology” as was shown in the Sections 1.-5. above, it should be possible to replace the word “ontology” used in knowledge representation view by using more common words such as a “concept”, a “language”, and a “world”, or possible “worlds”. Usually, when we want to give an explication of some particular concept, we replace of a vague concept (the *explicandum*) by a more precise one (the *explicatum*), i.e. we are searching out new definitions for old concept. The new definitions should be superior to the old in clarity, exactness, and it should fit into systematic structure of concepts. Now, in this case, a ‘sect’ of information systems community is doing *exactly opposite*.

Moreover, if we are not going to follow the logical positivist view of ontology by simply rejecting all the ontological problems appearing in philosophy, cf. [25], a deviation from the standard philosophical use of the word “ontology” used by other scientists and in other disciplines creates communication problems and will estrange and isolate that ‘sect’ of information systems community. Also, a language-dependent “ontology” is constrained with the ontological commitment inherent in the structure of the used language, i.e. as Wittgenstein says it in *Tractatus Logico-Philosophicus* 5.6: “The limits of my language mean the limits of my world” [20].

Conclusion

In this paper the role and the proper place of ontology was considered in the context of information systems and conceptual modelling. Firstly, an information systems and a conceptual modelling were described *without* using the word “ontology”. Then two senses of ontology are presented: philosophical and knowledge representation views. In this connection both Gruber’s and Guarino’s definitions of “ontology” were scrutinized. An “ontology” for Guarino as well as for Gruber was seen to be a language-dependent, which takes in account of its explication both concepts and the world(s).

Unfortunately, the kind of terminological shifts concerning the word “ontology”, and practiced e.g. by Gruber and Guarino, have caused more confusion than clarification in the field of information systems and conceptual modelling. Accordingly, it should be better to use the word “ontology” in its traditional philosophical sense only, whereas our aim in the field of information systems and conceptual modelling is to reach a conceptual clarity.

---

3 The same hold also for The OWL Web Ontology Language, [26], as well as The Protégé Ontology Editor, [27].
References:


[27] The Protégé Ontology Editor and Knowledge Acquisition Systems, (available at http://protege.stanford.edu/)