

Introduction: What is Ontology for?

Katherine Munn

If you are reading this, then chances are you are a philosopher, an information scientist, or a natural scientist who uses automated information systems to store or manage data.

What these disciplines have in common is their goal of increasing our knowledge about the world, and improving the quality of the information we already have. Knowledge, when handled properly, is to a great extent cumulative. Once we have it, we can use it to secure a wider and deeper array of further knowledge, and also to correct the errors we make as we go along. In this way, knowledge contributes to its own expansion and refinement. But this is only possible if what we know is recorded in such a way that it can quickly and easily be retrieved, and understood, by those who need it. This book is a collaborative effort by philosophers and information scientists to show how our methods of doing these things can be improved. This introduction aims, in a non-technical fashion, to present the issues arising at the junction of philosophical ontology and information science, in the hope of providing a framework for understanding the essays included in the volume.

Imagine a brilliant scientist who solves a major theoretical problem. In one scenario he scribbles his theory on a beer mat, sharing it only with his drinking companions. In this scenario, very few scientists will have the ability to incorporate this discovery into their research. Even were they to find out that the solution exists, they may not have the resources, time, or patience to track it down. In another scenario our scientist publishes his solution in a widely read journal, but has written it in such a sloppy and meandering way that virtually no one can decipher it without expending prohibitive amounts of effort. In this scenario, more scientists will have access to his discovery, and may even dimly recognize it as the truth, but may only understand it imperfectly. No matter how brilliant our scientist is, or how intricately he himself understands his discovery, if he fails to convey it to the scientific community in such a way that they have ready access to it and can understand it, unfortunately that community will not benefit from what he has discovered. The moral of this story is that the means by which knowledge is conveyed are every bit as important as that knowledge itself.

The authors' goal in producing this book has been to show how philosophy and information science can learn from one another, so as to

create better methodologies for recording and organizing our knowledge about the world. Our interest lies in the representation of this knowledge by *automated* information systems such as computerized terminologies and taxonomies, electronic databases, and other knowledge representation systems. Today's automation of knowledge representation presents challenges of a nature entirely different from any faced by researchers, librarians or archivists of the pre-computer age.

Before discussing the unique challenges posed by automated systems for storing knowledge, we must say a few brief words about the term 'knowledge'. We are not using this term in a sense corresponding to most philosophical theories. What these theories have in common is the requirement that, in order for a belief or a state of mind to count as knowledge, it must connect the person to the truth. That is, a belief or a state of mind counts as knowledge only if its representational content corresponds with the way the world is. Most philosophical theories add the condition that this correspondence must be non-accidental: there must be a causal relation between the belief and its being the case; the person must base the belief on a certain kind of evidence or justification, and so forth (pick your theory).

The sense of 'knowledge' used in information science is more relaxed. Terms such as 'knowledge engineering' and 'knowledge management' do not refer to knowledge in the sense of a body of beliefs that are *apodictically true*, but of a body of beliefs which the scientific community has good reason to believe are true and thus treats in every respect as if they are true. Most researchers recognize that some of these highly justified beliefs are not, in fact, knowledge in the strict sense, since further scientific development could show them to be false. Recognizing this is part of what drives research forward; for part of the goal of research is to cause the number of false beliefs to decrease and the number and nuance of true beliefs to increase. The information stored in automated systems constitutes knowledge in the sense of *beliefs which we have every reason to believe are true, but to which we will not adhere dogmatically should we obtain overruling reasons to believe otherwise*. (We will often use 'information' in the same sense as 'knowledge'.) This approach, called *realist fallibilism*, combines a healthy intellectual humility with the conviction that humans can take measures to procure true beliefs about the world.

So much for 'knowledge'. What does it mean to *store* or *represent* knowledge? (We will use these terms interchangeably.) Say that you have a

bit of knowledge, i.e., a belief that meets all the requirements for knowledge. To store or represent it is to put it into a form in which it can be retained and communicated within a community. Knowledge has been stored in such forms as words, hieroglyphs, mnemonics, graphs, oral tradition, and cave scratching. In all of these forms, knowledge can be communicated, passed on, or otherwise conveyed, from one human being to another.

Automated information systems pose unprecedented challenges to the task of storing knowledge. In the same way that knowledge is represented on the pages of a book by one person and read by another, it is entered into an automated system by one person and retrieved by another. But whereas the book can convey the knowledge to the reader in the same form in which the writer recorded it, automated information systems must store knowledge in forms that can be processed by non-human agents. For computers cannot read or understand words or pictures, so as to answer researchers' queries in the way that the researchers would pose them, or to record their findings as researchers would. Computers must be programmed using explicit codes and formulas; hence, the quality of the information contained in information systems is only as high as the quality of these codes and formulas.

Automated information systems present unique opportunities for representing knowledge, since they have the capacity to handle enormous quantities of it. The right technology enables us to record, obtain, and share information with greater speed and efficiency than ever before, and to synthesize disparate items of information in order to draw new conclusions. There are different sorts of ways in which information systems store knowledge. There are databases designed for storing *particular* knowledge pertaining to, for example, specific experimental results, specific patients treated at a given hospital during a given time period, or specific data corresponding to particular clinical trials. Electronic health record (EHR) systems, used by hospitals to record data about individual patients, are examples of databases which store such particular knowledge. There are also systems designed for storing *general* knowledge. General knowledge includes the sorts of statements found in textbooks, which abstract from particular cases (such as *this* patient's case of pneumonia) and pertain, instead, to the traits which most of those particular cases have in common (such as lung infection, chill, and cough). Systems designed to store general knowledge include controlled vocabularies, taxonomies, terminologies, and so forth. Examples of these

include the Gene Ontology, the Foundational Model of Anatomy, and the Unified Medical Language System Semantic Network.

Ideally, these two types of system will play complementary roles in research. Databases and other systems for storing particular information should be able to provide empirical data for testing general theories, and the general information contained in controlled vocabularies and their ilk should, in turn, provide sources of reference for empirical researchers and clinicians. How better, for example, to form and test a theory about pneumonia than by culling the clinical records of every hospital which has recorded cases of it? How better to prepare for a possible epidemic than by linking the electronic record systems of every hospital in the country to a centralized source, and then programming that source to automatically tag any possibly dangerous trends?

But in order for these goals to be realized, automated information systems must be able to share information. If this is to be possible, every system has to represent this information in the same way. For any automated information system to serve as a repository for the information gathered by researchers, it must be pre-programmed in a way that enables it to accommodate this information. This means that, for each type of input an information system might receive, it must have a category corresponding to that type. Therefore, an automated information system must have a categorial structure readymade for slotting each bit of information programmed into it under the appropriate heading. That structure, ideally, will match the structure of other information systems, to facilitate the sharing of information among them. But if this is to be possible, there must be one categorial structure that is common to all information systems. What should that structure look like?

There are several possible approaches to creating category systems for representing information about the world. One approach, which Smith calls the *term orientation* (see Chapter 4), is based on the observation that researchers often communicate their findings in the form of sentences. What better way to create a category system than to base it on the meanings of the words in those sentences? One problem with this approach is that the meaning of a word often does not remain constant; it may change from context to context, as well as over the course of time. Another problem is that natural language cannot be guaranteed to contain a word which encompasses precisely the meaning one wants to express, especially in scientific disciplines that are constantly making discoveries for which there are not yet established words. Another approach, which is standardly

referred to as the *concept orientation*, attempts to get around these difficulties by substituting words with concepts, seen (roughly) as hypostatizations of the meanings of words into mental entities. In other words, a concept is a word whose meaning has been fixed forever in virtue of being attached to a special kind of abstract thing. Thus, even if some slippage occurs between a word and its original meaning, that meaning will always have a concept to which it adheres. One simple problem with this approach (Smith provides a litany) is that it goes to great lengths to posit a layer of reality – that of concepts – for theoretical purposes only. This raises the question why the structure of the world itself should not be used as a guide to creating categories, an approach known as *realism*. After all, our knowledge is *about* the world, not about concepts.

A major contention against realism is that reality is just too massive, diffuse, or limitless, for human understanding to grasp. There are far more things in the world, and far more kinds of things, than any one person can think or know about, even over the course of a lifetime. Ask one hundred people what the most basic underlying categories of the world are, and you will likely get one hundred different answers. Even scientific disciplines, which reflect not the understanding of one person but of successive groups of people with similar goals and methods, can produce no more than a perspective on one specific portion of reality, to the exclusion of the rest. The object of their study is limited to a specific *domain* of reality, such as the domain of living things for biology or the domain of interstellar-objects for astronomy. Human understanding cannot, either individually or collectively, grasp reality as it is in its entirety; hence, the conceptualist does not expect to be able to represent reality in the categories of automated information systems.

The realist response developed in this volume (particularly Chapters 1, 3, 4, 6, and 7) is this: we can and should understand the existence of multiple perspectives not as a hindrance to our ability to grasp that reality as it is, but as a means by which we can obtain a deeper understanding of it. For, from the fact that there are multiple perspectives on reality alone, it does not follow that none – or only one – of these perspectives is *veridical*, i.e., represents some aspect of reality *as it truly is*.

A perspective is merely the result of someone's coming to cognitive grips with the world. Precisely because reality is so multi-faceted, we are forced to filter out some aspects of it from our attention which are less relevant to our purposes than others. Some of these processes of selection are performed deliberately and methodically. For example, biologists set

into relief the domain of living things, in order to focus their study on traits shared by them which non-living things do not have. Forest rangers set into relief the domain of a specific geographical area and certain specific features, such as marked trails and streams, which they represent in maps for the purposes of navigation. Often, especially among scientists, the purpose of roping off a particular domain is simply to gain understanding of what the entities within it have in common, and of what makes them different from entities in other domains.

The selection of a particular perspective is an act of cognitively *partitioning* the world: drawing a mental division between those things upon which we are focusing and those which fall outside our domain of interest. (Chapter 6 develops a theory of how we partition the world.) Take as an example Herbert, who is a frog. Let us imagine that Herbert is a domain of study unto himself. We thereby cognitively divide the world into two domains: Herbert, and everything else.

Given a partitioning of the world into domains, it becomes possible to create sub-partitions within those domains. Herbert happens to be a frog, in addition to being composed of molecules. Each of these features yields a unique perspective from which Herbert can be apprehended: the coarse-grained level of Herbert as a whole single unit, and the fine-grained level of his molecules. Most of us think of Herbert as a single unit because it is as such that we apprehend him in his terrarium. Although we may know that he is composed of molecules, his molecules are not relevant to our apprehension of him, and so we filter them out. A molecular biologist, on the other hand, may think more about Herbert's molecules than about Herbert as a whole, even though he is aware that those molecules constitute a whole frog. There is only one Herbert that we and the molecular biologist apprehend, but, depending upon our interests and our focus, we may each apprehend him from different *granular perspectives*.

Recognizing that there are multiple veridical perspectives on reality is *not* equivalent to endorsing relativism, the view that *all* perspectives are veridical. Here are two examples of non-veridical perspectives on Herbert: one which views him as a composite of the four complementary elements earth, air, fire, and water; another which views him as an aggregate of cells joined by an aberrant metaphysical link to the soul of Napoleon. The existence of multiple perspectives does not imply that we are unable to grasp reality as it is, and the fact that it is possible to obtain deeper understanding of reality *through* those perspectives does not imply that all perspectives are veridical representations of reality.

This is not to suggest that it is always easy to distinguish veridical perspectives from non-veridical ones. In fact, it is this difficulty which forces responsible ontologists and knowledge engineers to temper their realism with a dose of fallibilism. One of the main ways to determine the likelihood of a perspective's being veridical is to assess its explanatory power, that is, the breadth and depth of the explanations it can offer of the way the world works. The four-element perspective on Herbert seemed plausible to certain people at a certain point in history, precisely because it offered a means of explaining the causal forces governing the world. It seems less plausible now because better means of explanation have been developed.

Each automated information system strives to represent a veridical perspective on that partition of reality about which it stores knowledge. As we have seen, there are features intrinsic to such systems which render them better or worse for fulfilling this goal. A system which is programmed with a structure that corresponds closely to the structure of the granular partition itself is more likely to be veridical; think of the four-element perspective versus the molecular one. An information system with the categories 'earth', 'air', 'fire', and 'water' is less likely to serve as basis for an accurate categorization of Herbert's various components than is a system with such categories as 'cell', 'molecule', and 'organ'.

The best kinds of categories are *natural* in the sense that they bring genuine similarities and differences existing in reality to the forefront (this view is developed in Chapters 7 and 8). Natural category divisions tell us something about how the underlying reality truly is. Thus, it is more likely that knowledge of such naturally existing categories will put us in a position to construct systematic representations of that domain which have some degree of predictive power. If we can predict the way in which entities in a domain will behave under certain conditions, we are better able to understand that domain, interact with it, and gain more knowledge about it.

Hence the realist, who believes that it is possible for humans to obtain knowledge about the world, seeks to find out, as best he can, what the natural categories of reality are. His goal as a knowledge engineer is to create an information system that is structured in a way that mirrors those categories. Such a system will be prepared to receive information about as wide an array of entities as possible. Then, it should represent information by tagging each piece of information as being about something that has certain traits which make that thing naturally distinct from other entities.

Now, there is at least one natural category into which every entity falls: the category of existing things. It follows that there is at least one perspective from which all of reality is visible, one partition in which every entity naturally belongs: the partition of existing things. This partition is admittedly large-grained in the extreme; it does not provide us with more than a very general insight into the traits of the entities it encompasses. But it does provide us with insight into one crucial trait, existence, which they all have in common. It is this partition which constitutes the traditional domain of *ontology*.

Ontology in the most general sense is the study of the traits which all existing things have insofar as they exist. (This is an admittedly airy definition of an abstract notion; see Chapter 2 for elaboration). It is significant that the philosophical term ‘ontology’ has been adopted by the information-science community to refer to an automated representation (taxonomy, controlled vocabulary) of a given domain (a point developed in Chapter 1). We will sometimes use the term ‘ontology’ in this sense, in addition to using the philosophical sense expounded in Chapter 2.

Since there is one trait, existence, which all entities in reality have in common at the most general level, it is reasonable to suppose that there are other traits which some entities have in common at more specific levels. This supposition conforms to our common-sense assumption that some entities are more alike than others. If this is correct, it would suggest that our ability to understand something about reality in its entirety does not stop at the most general level, but continues downward into more specific levels. The challenge for the realist is to devise a means to discern the categorial subdivisions further down the line; this challenge is taken up in Chapter 9.

Clearly, an upper-level system of categorization encompassing all entities would be an enormous step toward the goal of optimal knowledge representation. If all information systems were equipped with the same upper-level category system (sometimes called a *domain-independent formal ontology*), and if this category system did exhaust the most general categories in reality, then it would be possible to share information among systems with unprecedented speed, efficiency, and consistency. The contributions in this book are aimed at this long-term, but worthwhile, goal. Although the methods developed here are intended to be applicable to any domain, we have chosen to limit our focus primarily to the domains of biology and medicine. The reason is that there are particularly tangible benefits for the knowledge representation systems in these domains.

Accordingly, in ‘Bioinformatics and Philosophy’ (Chapter 1), philosopher Barry Smith and geneticist Bert Klagges make a case for the use of applied ontology in the management of biological knowledge. They argue that biological knowledge-management systems lack robust theories of basic notions such as kind, species, part, whole, function, process, environment, system, and so on. They prescribe the use of the rigorous methods of philosophical ontology for rendering these systems as effective as possible. Such methods, developed precisely for the purpose of obtaining and representing knowledge about the world, have a more than two thousand year-old history in knowledge management.

In ‘What is Formal Ontology?’ (Chapter 2) Boris Hennig brings that most general, abstract domain of existing things down to earth. His goal is to help us understand what the more specific categories dealt with in this book are specifications of. The historical and philosophical background he provides will enable us to view formal ontology afresh in the present context of knowledge management. That context is illuminated in Pierre Grenon’s ‘A Primer on Knowledge Management and Ontological Engineering’ (Chapter 3). Grenon draws upon non-technological examples for two purposes: first, to explain the task of knowledge management to non-information scientists; second, to highlight the reasonableness of the view that knowledge management is about representing reality. He provides insight into the task of the knowledge engineer, who is promoted to the post of *ontological engineer* when he uses rigorous ontological methods to systematize the information with which he deals. Finally, Grenon describes some current (worrying) trends in the knowledge-management field, for which he prescribes a realist ontological approach as an antidote.

Some of these trends are elaborated upon in Barry Smith’s ‘New Desiderata for Biomedical Terminologies’ (Chapter 4). Smith chronicles the development of the concept orientation in knowledge management, offering a host of arguments against it and in favor of the realist orientation. In ‘The Benefits of Realism: *A Realist Logic with Applications*’ (Chapter 5) Smith goes on to demonstrate the problem-solving potential of a realist orientation. He does so by developing a methodology for linking sources of particular knowledge (such as databases) with sources of general knowledge (such as terminologies) in order to render them interoperable. This would dramatically improve the speed and efficiency of the information-gathering process as well as the quality of the information garnered. Implementing his methodology would require a global switch to

the realist orientation in knowledge management systems. Arduous as such a switch would be, his example shows the massive benefits that it would proffer.

If we are to reconstruct existing knowledge management systems to reflect a realist orientation, we will need a theoretical blueprint to guide us. We must start by formalizing the most basic commitment of the realist orientation, *realist persepectivalism*, which is the view that we can obtain knowledge of reality itself by means of a multiplicity of veridical granular partitions. Bittner and Smith (Chapter 6) provide a formal theory of granular partitions for configuring knowledge management systems to accommodate the realist orientation. Only such a theory, they claim, can provide the foundation upon which to build knowledge management systems which have the potential to be interoperable, even though they deal with different domains of reality.

How do we build up an information system that succeeds at classifying the entities in a given domain on the foundation of a theory of granular partitions? In 'Classifications' (Chapter 7), Ludger Jansen provides eight criteria for constructing a good classification system, complete with real examples from a widely used information system, the National Cancer Institute Thesaurus (NCIT), which fails to meet them. Nonetheless, he points out, there are numerous practical limitations which an ontological engineer must take into account when constructing a realist ontology of his domain. Since a classification system is, to some extent, a model of reality, the more limited the knowledge engineer's resources (temporal, monetary, technological, and so forth), the greater his system must abstract from the reality it is supposed to represent. But the existence of such practical limitations does not require us to abandon the goal of representing reality. Jansen recommends meeting practical needs with accuracy to reality by distinguishing between two types of ontologies with distinct purposes. The purpose of *reference ontologies* is to represent the complete state of current research concerning a given domain as accurately as possible. Alternatively, the purpose of *application ontologies*, such as particular computer programs, should be to fit the most relevant aspects of that information in an application designed with certain practical limitations in mind. Reference ontologies should serve as the basis for creating application ontologies. This way, accuracy to reality can stand side by side with utility without either one needing to be sacrificed. Further, application ontologies that are based on the same reference ontologies will be more

easily interoperable with each other than application ontologies based on entirely different frameworks.

In ‘Categories: *The Top-Level Ontology*’ (Chapter 8), Jansen applies the criteria for good classification to the question of what the uppermost categories of a reference ontology should be. Once we move below the most general category, ‘being’, what are the general categories into which all existing things can be exhaustively classified? Jansen answers this question by drawing upon the work of that most famous philosopher of categories, Aristotle. He provides examples of suggested upper-level ontologies which are currently in use, the Suggested Upper Merged Ontology (SUMO) and the Sowa Diamond, and argues that they are inferior to Aristotle’s upper-level categories. He then presents the upper-level category system Basic Formal Ontology (BFO), which was constructed under the influence of the Aristotelian table of categories, and makes the case for using BFO as the standard upper-level category system for reference ontologies.

Chapter 9 offers an example of the way in which Jansen’s considerations can be applied in one sort of theory that underpins the biomedical domain: the theory of the classification of living beings. On the basis of both philosophical and practical considerations, Heuer and Hennig justify the structure of the traditional, Linnaean, system of biological classification. Then they discuss certain formal principles governing the development of taxonomies in general, and show how classification in different domains must reflect the unique ontological aspects of the entities in each domain. They use these considerations to show that the traditional system of biological classification is also the most natural one, and thereby also the best.

Knowing how existing things are to be divided into categories is the first step in creating a reference ontology suitable for representing reality. But this is not enough. In addition to knowing what kinds of entities there are, we must know what kinds of *relations* they enter into with each other. We learn about the kinds of entities in reality by examining instances of these entities themselves. In ‘Ontological Relations’ (Chapter 10), Ulf Schwarz and Barry Smith argue that this is also the way to learn about the kinds of relations which obtain between these kinds of entities: we must examine the particular relations in which particular entities engage. They endorse the efforts of a group of leading ontological engineers, the Open Biomedical Ontologies (OBO) Consortium, to delineate the kinds of relations obtaining between the most general kinds of entities.

In Chapter 11, Ingvar Johansson offers a detailed treatment of one of the relations discussed in Chapter 10, the so-called *is_a* or subtype relation, which plays a particularly prominent role in information science. Johansson argues that there are good reasons to distinguish between four relations often confused when *is_a* relations are intended: genus-subsumption, determinable-subsumption, specification, and specialization. He shows that these relations behave differently in relation to definitions and so-called inheritance requirements. From the perspective predominant in this book, classifications should be marked by the feature of single inheritance: each species type in a classification should have a single parent-type or genus. The distinction between *single inheritance* and *multiple inheritance* is important both in information science ontologies and in some programming languages. Johansson argues that single inheritance is a good thing in subsumption hierarchies and is inevitable in pure specifications, but that multiple inheritance is often acceptable when *is_a* graphs are constructed to represent relations of specialization and in graphs that combine different kinds of *is_a* relations.

Many relations obtain between *continuant* entities; that is, entities, such as chairs and organisms, which maintain their identity through time. But reality also consists of processes in which continuant entities participate, which form a different category of entity, namely, *occurrent* entities. Just like continuants, occurrents can – and must – be classified by any information system which seeks a full representation of reality. For, just as there are continuants such as diseases, so there are the occurrents that are referred to in medicine as disease courses or disease histories. Hennig's 'Occurrents' (Chapter 12) develops an ontology, or classification, of occurrent entities. He distinguishes between processes, which have what he calls an internal temporal structure, and other temporally extended occurrents, which do not. Further, he notes that certain important differences must be taken into account between types of occurrents and their instances. He argues that particular occurrents may instantiate more than one type at the same time, and that instances of certain occurrents are necessarily incomplete as long as they occur. By pointing out these and other important ways in which occurrents differ from continuants, Hennig's work shows the urgency of the need for information systems to obtain clarity in their upper-level categories.

Finally, in Chapter 13, Johansson takes a wide-lens view of the junction of philosophy, ontology, and bioinformatics. He observes that some bioinformaticians, who work with terms and concepts, are reluctant to

believe that it is possible to have knowledge of mind-independent reality in the biological domain. He argues that there is no good reason for this tendency, and that it is even potentially harmful. For, at the end of the day, bioinformaticians cannot completely disregard the question as to whether the terms and concepts of their discipline refer to real entities. In the first part of the chapter, Johansson clarifies three different positions in the philosophy of science with which it would be fruitful for bioinformaticians to become familiar, defending one of them: Karl Popper's epistemological realism. In the second part, he discusses the distinction (necessary for epistemological realism) between the *use* and *mention* of terms and concepts, showing the importance of this distinction for bioinformatics.

This volume does not claim to have the final say in the new discipline of applied ontology. The main reason is that the ideas it presents are still being developed. Our hope is that we have made a case for the urgency of applying rigorous philosophical methods to the efforts of information scientists to represent reality. That urgency stems from the vast potential which such application can have for rendering information systems interoperable, efficient, and well-honed tools for the increasingly sophisticated needs of anyone whose life may be affected by scientific research – that is to say, of everyone. What the authors of this volume are working toward is a world in which information systems enable knowledge to be stored and represented in ways that do justice to the complexity of that information itself, and of the reality which it represents.