# Ontology: Not Just for Philosophers Anymore

**Robert Arp**

## Abstract

*Following a brief discussion of informatics and of the problems it faces, the distinction is introduced between philosophical ontology, domain ontology, and formal ontology. After giving examples of the kinds of logical and conceptual problems discoverable in domain ontologies that have hindered progress in informatics, it is then shown how a formal, ontological approach inspired by ideas and methods of philosophy can assist in ensuring the synchronized development of domain ontologies in such a way as to promote the optimal retrieval and dissemination of information. Next, some concrete steps are offered that one may take in constructing a domain ontology that is rationally coherent, optimally computable, and interoperable with other domain ontologies. Finally, the first steps in developing a philosopher's résumé domain ontology is offered that will assist in the sharing and retrieval of basic information that one usually finds on a philosopher's résumé or curriculum vita, using Heidegger as the primary example.*

**Keywords:** Aristotle, Basic Formal Ontology, domain ontology, formal ontology, Heidegger, Husserl, informatics, information, interoperability, logic, ontology, philosopher's résumé domain ontology, silo effect

## Information Overload

Now, as never before in history, the amount of data and information that is being made available regarding various sciences, domains, and disciplines is quite literally overwhelming. *Informatics* is the science associated with the collection, categorization, management, storage, processing, retrieval, and dissemination of information—principally, through the use of computers as well as computational and mathematical models—with the overall goal of improving retrieval and dissemination of information (Luenberger, 2006). Disciplines are increasingly developing their own informatics—for example, bioinformatics, medical informatics, legal informatics (Polanski & Kimmel, 2007)—reflecting the information overload that confronts researchers.

Thus, there is a very basic problem faced by all disciplines of simply collecting, classifying, and annotating information so that it may be made available through the World Wide Web. In a fairly recent *Scientific American* article, protagonists of the Semantic Web explicitly express the dream of a 'Great Encyclopedia' database that would assist in organizing information and, when queried, "give us a single, customized answer to a particular question without our having to search for information or pore through results" (Feigenbaum *et al.*, 2007, p.90). With the appropriate thinking and tools, this imagined repository has the potential to become reality.

The problem, then, is to organize this information in such a way that it can be efficiently accessed, shared, and used by human individuals. To assist in this organization, researchers have in recent years worked with computer scientists and programmers to set up what are known as *domain ontologies* in their fields of study. What exactly is a domain ontology?

## Domain Ontology

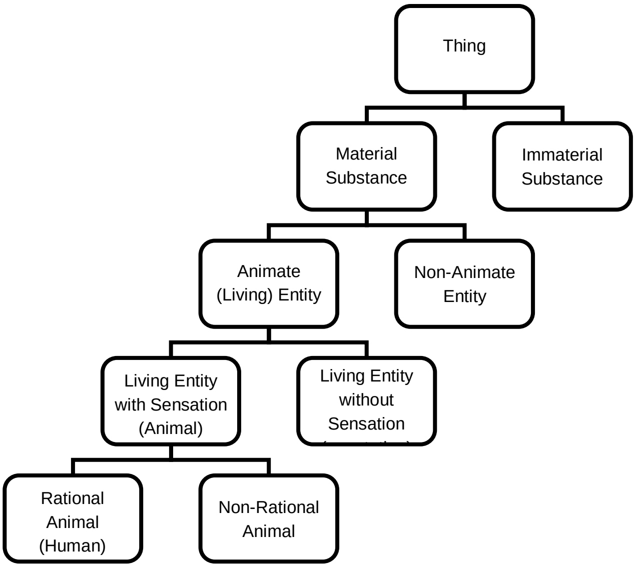
A *domain* is an area, sphere, or delineated portion of reality which humans seek to know, understand, and explain (possibly predict, manipulate, and control, too) as fully as is possible through the development of a subject matter, field, science, or discipline concerning that area. Examples include all of the various subjects investigated at a typical university, including medicine, engineering, law, economics, philosophy, psychology, and the like, complete with their respective, subsumed subject matters.

Philosophers can be heard making remarks such as the following:

- I don't see how one can fit wholesale evolution and a creating god into one's ontology without contradiction.
- There is often tension between the realist and antirealist ontological approaches to universals.
- Their work rested on an ontological presupposition according to which sense data formed the basic furniture of reality.

Here, the word 'ontology' is used in the traditional philosophical sense, referring to the branch of metaphysics that studies the nature of existence. From this *philosophical* perspective, ontology seeks to provide a definitive and exhaustive classification of entities and relationships in all domains or spheres of being, along the lines of what Porphyry attempted with his now famous Porphyrian Tree (Figure 1). The tree is a kind of taxonomy, a graph-theoretic representational artifact that is organized by hierarchical relations with leaves or nodes (representing *types*, *universals*, or *classes*) and branches or edges (representing the subtype relation).

*Figure 1: The Porphyrian Tree*



We are all naturally philosophical ontologists of one sort or another, since we all of us form systems of classification as we try to understand, navigate, control, and predict the complex workings of this universe. For example, we sort things into genus/species hierarchical relationships of greater and lesser degrees of complexity.

Now, consider these claims that might sound foreign to people in philosophical circles:

- I'm working on an ontology for MRI tests.
- The Gene Ontology has data on that HOX gene.

Related to this philosophical sense, for the past twenty years or so 'ontology' also has come to be understood as a structured, taxonomical representation of the entities and relations existing within a particular domain of reality such as geography, ecology, law, biology, medicine, or philosophy (Gruber, 1993; Smith, 2003; Arp, 2009). Domain ontologies, thus, are contrasted with ontology in the philosophical sense, which has all of reality as its subject matter.

A domain ontology, too, is a graph-theoretical representation, comprising a backbone taxonomic tree whose nodes represent types of entities in reality. These nodes are connected by edges representing principally the *is_a* subtype relation, but also supplemented by other edges representing binary relations such as *part_of*, *preceded_by*, *has_participant*, *inheres_in*, and other relations holding between these types of entities. To take a few common examples of assertions found in domain ontologies linking types together by means of such relations:

- cell nucleus (entity) *located_in* (relation) cell (entity),
- X-ray test *has_participant* patient,
- verdict *preceded_by* court trial,
- petroleum jelly *transformation_of* petroleum,

Further, the domain ontology contains properties and axioms that are designed to enable algorithmic reasoning on the basis of these relationships, so that new information about the underlying instances that comprise the domain of study might be inferred. For example, properties associated with *is_a* and *part_of* are transitive, enabling inferences such as:

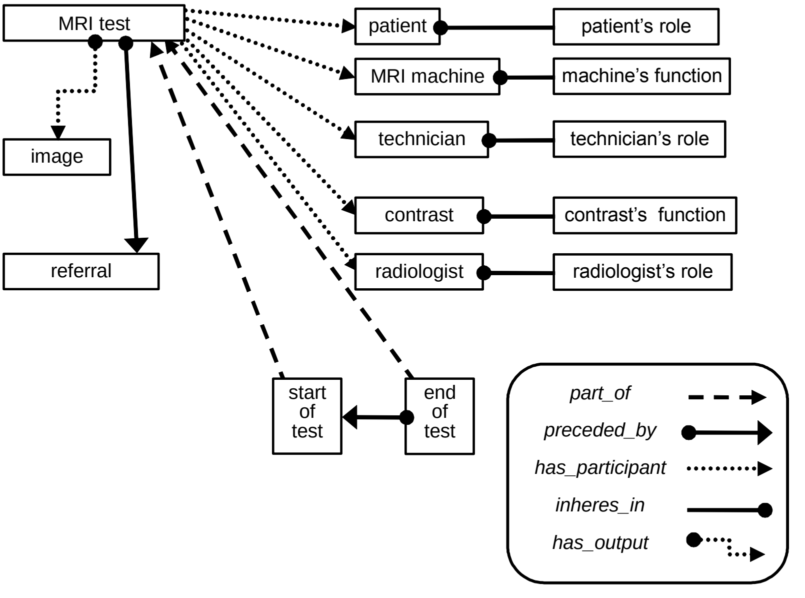
- brain *part_of* nervous system, and
- nervous system *part_of* body, therefore
    - brain *part_of* body.

- West Texas Intermediate petroleum *is_a* petroleum, and
- petroleum *is_a* flammable liquid, therefore
    - West Texas Intermediate petroleum *is_a* flammable liquid.

- flask's function *is_a* artifactual function, and
- artifactual function *is_a* function, therefore
    - flask's function *is_a* function.

Figure 2 represents part of an ontology devoted to the domain of MRI tests in the radiological sciences (Arp *et al.*, 2008), while Figure 3 represents part of a domain ontology devoted to cells (also see, for example, Arp & Smith, 2008).[1]

*Figure 2: The Beginnings of an Ontology for MRI Tests*

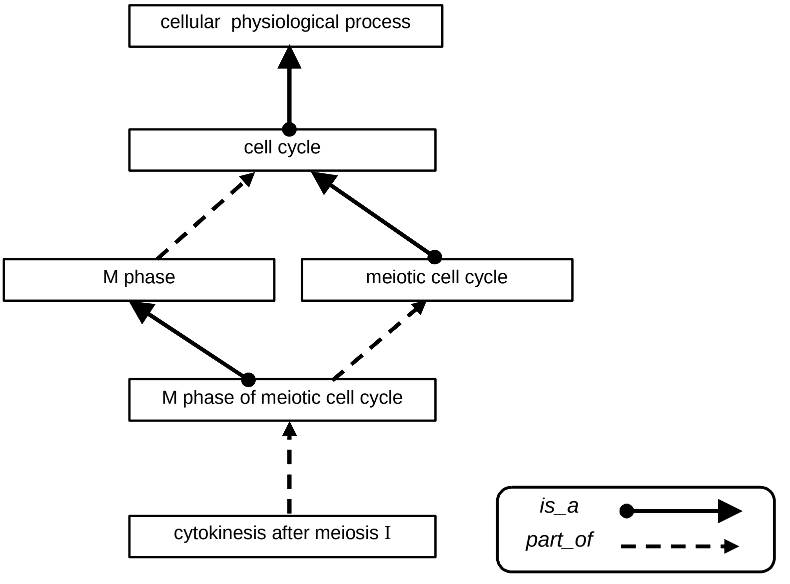



Other examples of domain ontologies currently being utilized by researchers around the world include the *Gene Ontology* (GO), *Foundational Model of Anatomy* (FMA), *Ontology for Biomedical Investigations* (OBI), *Protein Ontology* (PO), and many more biomedical ontologies available through the *Open Biomedical Ontologies* (OBO) Foundry at: http://obofoundry.org/. There are even domain ontologies being developed that are devoted to the philosophical disciplines through the work of researchers such as Kim, Choi, Shin, & Kim (2007).

In line with the informatician's goal of improving retrieval and dissemination of information, the purpose of a domain ontology is to make the information in the corresponding discipline more easily searchable by human beings and more efficiently and reliably processable by computers. Ontologies may in addition be designed to ensure that the different bodies of information collected by different researchers in the same domain should all be represented in the same way, which assists interoperability and shareability of that information. For this, however, distinct domain ontologies must be developed in tandem on the basis of some overarching organization and commonly accepted set of principles.

[1] Figure 3 is adapted from: http:// www.yeastgenome.org/help/GO.html

*Figure 3: Part of a Cell Domain Ontology*




## Problems for Domain Ontologies

To bring about such coordination, however, is very difficult. Currently, the terminologies used in conveying and organizing information in distinct domains are developed in ad hoc ways; often, terminologies and database schemas fall short of being interoperable even when prepared by researchers from the same departments, groups, or labs. Researchers in distinct disciplines speak different languages, use different terminologies, and format the results of their research in different ways, and these problems are inherited by the ontologies developed to support their work. The result is a *silo effect*: data and information are isolated in multiple, incompatible silos, and shareability and reusability is greatly limited. This result contradicts what Domingue and Motta (1999) rightly have claimed is one of the fundamental roles of an ontology, namely, "to support knowledge sharing and reuse" (p.104).

There are many factors that contribute to the silo effect. Errors abound in domain ontologies, often errors of a sort that would be standardly handled in an introductory logic course. Since the formal languages and associated reasoning systems available for use with ontologies are still fairly limited in their capacities (see, for example, Luger, 2008), the use of ontologies to bring about regimentation of the available information requires a maximum amount of clarity and precision at each step in the process of constructing a domain ontology.

This is where philosophers can make a contribution to the advance of ontology in informatics by pointing out the pitfalls of poor representation and reasoning that hamper information accessibility and dissemination. Ontology may not be just for philosophers anymore, but it turns out that many of the basic ideas and methodologies of philosophy can contribute significantly to the betterment of informatics and its fundamental goals (for example, Smith *et al*., 2006; Ceusters *et al*., 2004; Vizenor *et al*., 2004; Spear, 2006; Arp, 2008). The following sections contain examples of the types of conceptual and logical mistakes discoverable in domain databases. The examples come from the biology and medical domains, since these are the most developed, widely-used, and widely-accessible ontologies being put to use; however, one can find similar problems in other domains and disciplines.

**Simply Getting the Facts Wrong**

A problem encountered in computational repositories has to do with misinformation. For example, The Systematized Nomenclature of Medicine (SNOMED, http://www.snomed.org/) once asserted:

- both testes *is_a* testes.

SNOMED now has:

- both testes *is_a* structure of right testes, and
- amputation of toe *is_a* amputation of foot.

Also, the Biomedical Research Integrated Domain Group (BRIDG) has defined:

- 'animal' as "a non-person living entity" (http://www.bridgproject.org/);

while the Health Level 7 (HL7) organization has defined:

- 'living subject' as "A subtype of Entity representing an organism or complex animal, alive or not" (http://www.hl7.org/).

**Examples Instead of Definitions**

To give a definition is to explicate, clearly and coherently, the essential distinguishing feature or features of the thing which make it be what it is. One of the first lessons we learn when trying to grasp the meaning of a term is to provide an *example*. However, oftentimes people will utilize an example and mistakenly think it is the same thing as the definition of a term, or actually is the term, idea, or concept itself. For example, the BRIDG definition of 'adverse event' includes:

- toxic reaction,
- untoward occurrence in a subject administered a pharmaceutical product, as well as
- an unfavourable and unintended reaction, symptom, syndrome, or disease encountered by a subject on a clinical trial.

All of these would appear to be examples of types of adverse events. The definition does not provide any statement of necessary and sufficient conditions, and so we do not know what it is about these examples that would make them all "adverse events" (see Ceusters *et al*., 2008).

**Lack of Clear and Coherent Definitions**

It is not enough to offer a definition; we must offer one that is clear to its users, free from counter-examples, and provides both necessary and sufficient conditions. The attempt must be made to define terms, and some people have done an obviously poor job. For example, the *National Cancer Institute Thesaurus* (NCIT) has defined 'disease progression' simultaneously as:

cancer that continues to grow and spread, and
increase in size of tumor, and
1) the worsening of a disease over time.

So, which is it? Further, one will notice that: again, (1) and (2) are mere examples, instead of definitions; while (3) gives us, at best, a definition of 'progression' and says nothing about the definition of 'disease'.[2]

Also, in SNOMED we find:

- European *is_a* ethnic group,
- Other European in New Zealand is_a ethnic group,
- Unapproved attribute is_a function,
- Genus Mycoplasma is_a Prokaryote, and
- Prisheksninsk pig breed *is_a* organism.

What could these mean? Officially, in all the above cases, *'is_a'* is to be interpreted as meaning "is a subtype of," so that every instance of the type or class designated by the first term must be an instance of the type or class designated by the second term. It is just not true, however, that every European is an ethnic group, or that every pig breed is an organism.

**Circular Definitions**

A definition is circular if the term to be defined (*definiendum*) occurs in the definition itself (*definiens*). For then the definition can yield no new information regarding the meaning or referent of the term in question. Circular definitions are of course very easy to formulate; hence, one often finds circular definitions advanced by researchers in the information ontology world.
Consider BRIDG's definition of:

- 'ingredient' as "a substance that acts as an ingredient within a product,"

or the *Biomedical Informatics Research Network's* (BIRNLex)[3] definition of:

- 'eyeball' as "the eyeball and its constituent parts."

**Perception/Conception vs. Reality Confusion**

It is a good rule of thumb, from the perspective of philosophical ontology, to distinguish between (a) what one perceives/conceives/knows to be the case and (b) what actually is the case irrespective of one's perceptions, conceptions, or knowledge. Yet, in the realm of information ontology, researchers often conflate or confuse the two.

Consider, again, BRIDG:

- Living subject =def. An object representing an organism.
- Class performed activity =def. The description of applying, dispensing, or giving agents or medications to subjects.
- Adverse event =def. An observation of a change in the state of a subject that is assessed as being untoward.

[2] see http://www.nci.nih.gov/cancerinfo/terminologyresources

[3] BIRNLex: http://xwiki.nbirn.net/xwiki/bin/view/+BIRN-OTF-Public/Home

- Objective result =def. An act of monitoring, recognizing and noting reproducible measurement.

It is, once again, an error of the type standardly addressed in introductory philosophy classes, to confuse a living subject, an activity, and an event—which are entities out there in reality—with mental representations, descriptions, observations, models, or documentations thereof.

### Use-Mention Confusion

It is a very common mistake in the construction of ontologies—and other kinds of representational artifacts—to confuse or conflate claims that are about the ontology itself with claims that refer to objects in reality. Consider this example from the BIRNLex:

- mouse =def. name for the species *Mus musculus*.

Further, in the *Medical Subject Headings* (MeSH) database, one can find

- National Socialism *is_a* MeSH Descriptor

which confuses National Socialism as an actual political movement with the term 'National socialism' as a MeSH Descriptor. Under "definition of function" MeSH provides the following:

- used with organs, tissues, and cells of unicellular and multicellular organisms for normal function. It is used also with biochemical substances, endogenously produced, for their physiologic role.[4]

This not only confuses the use and mention of function, but it also conflates function with role.

Figure 4 lists a few of the basic pitfalls of poor reasoning spoken about in these last few sections that hamper information accessibility and dissemination, pitfalls of which domain ontologists should be mindful. The reader may notice that one would find these pitfalls presented in a basic logic class, giving further credence to the claim that many of the basic lessons learned through philosophical ontology might contribute to the improvement of ontology in informatics and its fundamental goals of information accessibility and shareability.

*Figure 4: Some Basic Pitfalls of Poor Reasoning to Avoid in Domain Ontologies*

- Simply Getting the Facts Wrong
- Examples Instead of Definitions
- Lack of Clear and Coherent Definitions
- Circular Definitions
- Perception/Conception vs. Reality Confusion
- Use-Mention Confusion

[4] http://www.nlm.nih.gov/cgi/mesh/2008/MBcgi

## Formal Ontology

Domain ontologies as currently constituted have, as already noted, contributed to the problem of data silos. To help remedy this problem, a third kind of ontology—an upper-level or *formal* ontology—has emerged. Formal ontology is designed to assist in organizing domain ontologies in such a way as to make the data and information they are used to annotate interoperable.

Insofar as it concerns informatics, the goal of formal ontology is the calibration of the domain ontologies constructed in its terms so that they form one single, organized, interconnected, and interoperable repository.

The term 'formal ontology' was coined by Edmund Husserl (1900, 1901) in his *Logical Investigations*. As Smith & Smith (1995) note, 'formal' means applicable to "all domains of objects whatsoever… independent of the peculiarities of any given field of knowledge" (p.28). Nowadays, the word 'formal' is used in ontology contexts interchangeably with 'upper-level,' 'top-level,' or 'higher-level', and this is appropriate since 'formal ontology' refers to a discipline which assists in making possible communication between and among domain ontologies (envisioned as mid- or lower-level ontologies) by providing a common formal framework or ontological backbone. A formal ontology aims to give a common internal structure to domain ontologies by providing a common formal framework for the categorization of entities in a way that also supports logical reasoning about those entities. So, ideally and analogously, whereas a domain ontology assists in organizing the information of a particular domain, a formal ontology assists in organizing the information annotated through multiple domain ontologies in such a way as to make all of the latter interoperable and uniformly accessible to interested parties. Some examples of formal ontologies include:[5]

- Standard Upper Merged Ontology
- Descriptive Ontology for Linguistic and Cognitive Engineering (DOLCE)
- *Basic Formal Ontology* (BFO).

### Basic Formal Ontology (BFO)

An ontology that is increasingly being utilized by domain ontologists in the sciences is *Basic Formal Ontology* (BFO). BFO was conceived and developed by Barry Smith, Pierre Grenon, and others (Grenon & Smith, 2004), and is being tested by developers of domain ontologies in a variety of scientific domains, primarily in the area of biology and medicine. It can claim a number of advantages relative to SUMO and DOLCE, including:

1. it is very small, reflecting a deliberate aim not to compete with the domain ontologies developed by scientists themselves;
2. it is being created through a combination of philosophical expertise and scientific testing, involving major communities of investigators—for example within the *Ontology for Biomedical Investigations* (OBI) consortium: http://obi-ontology.org.

[5] SUMO: http://suo.ieee.org; DOLCE: http://www.loa-cnr.it; BFO: http://www.ifomis.org/bfo.

BFO divides entities into their most basic categories, starting with *continuants* (entities that *continue* or persist through time—for example, objects, qualities, and functions) and *occurrents* (entities which *occur* in time—for example, processes, processual contexts, and processual boundaries). These categories are linked together by means of various kinds of relationships, including *is_a*, *part_of*, *has_participant*, and others of the kind already mentioned in the discussion of domain ontologies above (Smith, *et al.*, 2005). Illustrations of BFO categorizations can be found in Figure 5, which includes assertions to the effect that:

- object *is_a* independent continuant,
- independent continuant is_a continuant,
- process is_a processual entity,
- processual entity *is_a* occurrent,

and so forth.

*Figure 5: The Continuant and Occurrent Categories of Basic Formal Ontology (BFO) Organized by the is_a Subtype Relation*

- BFO:entity
    - continuant
        - independent continuant
            - object
            - object boundary
            - object aggregate
            - fiat object part
            - site
        - dependent continuant
            - generically dependent continuant
            - specifically dependent continuant
            - quality
            - realizable entity
                - function
                - role
                - disposition
        - spatial region
            - zero-dimensional region
            - one-dimensional region
            - two-dimensional region
            - three-dimensional region

- BFO:entity
    - occurrent
        - processual entity
            - process
            - process boundary
            - process aggregate
            - fiat process part
            - processual context
        - spatiotemporal region
            - scattered spatiotemporal region
            - connected spatiotemporal region
                - spatiotemporal instant
                - spatiotemporal interval
        - temporal region
            - scattered temporal region
            - connected temporal region
            - temporal instant
            - temporal interval

The *Open Biomedical Ontologies* (OBO) Foundry initiative—which includes the developers of numerous domain ontologies such as the Gene Ontology, Ontology for Biomedical Investigations, and others mentioned earlier—has embraced BFO as the official formal ontology for the OBO Foundry (see http://obofoundry.org/; also Smith *et al.*, 2007). This is because BFO is seen as a way to ensure interoperability of the domain ontologies being created within the Foundry, since the developers of the latter employ the same upper-level ontological categories and relations. Results of BFO research have been incorporated into software applications produced by technology companies such as *Ontology Works* (http://www.ontologyworks.com/), *Ingenuity Systems* (http://www.ingenuity.com/), and *Computer Task Group*: (http://www.ctg.com/).

## Steps in Constructing a Domain Ontology

Philosophers, scientists, and other thinkers with different research foci and methodologies constantly interact and exchange information with one another. Nowadays, although walks to the library stacks to photocopy journal articles still occur, almost all of the interaction and exchange of scientific information happens electronically, involving use of the World Wide Web and of locally held data and information repositories where people have placed the results of their research. It is thus of vital importance that domain ontologies are calibrated so that information from one domain becomes inter-translatable and inter-interpretable with information from neighbouring domains. In what follows, a few steps in constructing a domain ontology are offered. For easy reference, the steps themselves can be found in Figure 6.

*Figure 6: First Steps in Constructing a Domain Ontology*

*Step 1: Determine the Purpose of the Domain Ontology.*
*Step 2: Provide an Explicit Statement of the Intended Subject Matter.*
*Step 3: Determine the Most Basic Universals and Relations, Clearly and Coherently Defining Them.*
*Step 4: Use Aristotelian Structure When Formulating Definitions.*
*Step 5: Put the Universal Terms in a Taxonomic Hierarchy, Adding the Relevant Relations.*
*Step 6: Regiment the Ontology to Ensure Logical Coherence and Accuracy of Information.*
*Step 7: Seek Interoperability by Using a Formal Ontology such as Basic Formal Ontology.*
*Step 8: Concretize the Ontology in a Computer Tractable Representational Artifact.*
*Step 9: Apply the Ontology, and Test the Results in a Computing Context.*

### Step 1: Determine the Purpose of the Domain Ontology.

Constructing a domain ontology is analogous to building a tool, so that the tool's intended usage needs to be established first. Thus, it is essential at the very beginning to ask the question: "What is the ultimate purpose of this domain ontology?"

This will include considering whether it is primarily intended to be a comprehensive representation of a given domain to serve as *reference* or *benchmark* (referred to as a *reference ontology*), or whether it is intended to be *applied* in order to accomplish certain more specific goals, such as data mining of the complete works of Aristotle (referred to as an *application ontology*).

### Step 2: Provide an Explicit Statement of the Intended Subject Matter.

Providing such a statement is essential for indicating what kinds of objects and relationships should be included in the domain ontology. Also, it forces researchers to concretize their thoughts about the purpose of the domain ontology (reference or application) in the form of a publicly available overview of the domain ontology. For example, the documentation for the *Foundational Model of Anatomy* reads, "The

FMA is strictly constrained to 'pure' anatomy, i.e., the structural organization of the body".[6]

**Step 3: Determine the Most Basic Universals and Relations, Clearly and Coherently Defining Them.**

An *instance* is a particular entity in the world, such as this particular frog on the dissection table here, the *Metropolitan Museum of Art* in New York City, or a particular performance of "Sweet Caroline" by Neil Diamond. On the other hand, a *universal* is a kind or type, like *frog*, *museum*, or *performance*. The goal of philosophy, science, or any body of knowledge is to make true statements about universals as well as devise generalizable laws, principles, or axioms (Lowe, 2006). Ontologies are representations not of particular instances, but rather of the universals which they instantiate. Once information regarding individual particulars is gathered, there is a natural process of sorting the information into categories on the basis of an identification of the universal features that the particulars share. This information is assembled on the one hand into scientific textbooks, and on the other hand into domain ontologies. An ontology is a representational artifact, comprising a taxonomy as proper part, whose representational units are intended to designate some combination of universals and the relations between them.

We want to know the definition of *cell membrane*, but we also want to know its relationship to other universals in the general schema of biology. Similarly, it is one thing to understand something about the universal *helium*, while it is another and much better thing to know how helium is related to the Periodic Table of the Elements and other aspects of reality. So, full knowledge of a given universal also requires understanding the relationships in which it stands to other universals, and conversely. To understand these relations we need to investigate the *instances* of the corresponding universals by performing scientific experiments.

Determining the universals and relations obtaining in a given domain is a matter of analyzing the subject matter, requiring the expertise of specialist scientists. Once a provisional determination of universals and relations occurs, then one can compile this into an initial list of terms with clear and associated definitions and with a tentative organization into categories provided by a top-level ontology such as Basic Formal Ontology. These terms should as far as possible reflect consensus usage in the corresponding discipline. The development of a domain ontology is an empirical endeavour, and thus the ontology itself never reaches a complete and finalized state.

**Step 4: Use Aristotelian Structure When Formulating Definitions.**

The Periodic Table, Linnean taxonomy, even the Porphyrian Tree itself, all owe their genesis to Aristotle's thinking. As we have seen, domain ontologies constructed thus far have manifested a less than adequate treatment of definitions. Given the central role in an ontology of the taxonomic (*'is_a'*) hierarchy, one strategy to resolve this problem is to adopt a rule according to which ontologies should use the Aristotelian structure when formulating definitions, which means that they should use definitions of the form "An A is a B that Cs", where B is the immediate *is_a* parent of A in the taxonomic hierarchy, and C is the defining characteristic of

---

[6] See http://sig.biostr.washington.edu/projects/fm/AboutFM.html

what picks out those Bs which are As (see Rosse & Mejino, 2003). Here are some examples:

- a human (A) is an animal (B) that is rational (C);
- a thyroid epithelial cell (A) is a cell (B) in the thyroid gland that secretes and produces the hormones thyroxine (T4) and tri-iodo-thyronine (T3);
- an odometer (A) is a device (B) that is used to indicate distance travelled by a vehicle (C).

This provides a consistent format for the representation of definitions that can be used regardless of the domain at issue; thus, it has a *formal* quality to it (recalling Husserl and the discussion above) and contributes to interoperability. Also, the definitions form natural, parent-child, taxonomic hierarchies based on the structure of the definitions alone. Further, this taxonomical structure makes computational inferences easier to perform, which is important for researchers using computational systems and the World Wide Web.

**Step 5: Put the Universal Terms in a Taxonomic Hierarchy, Adding the Relevant Relations.**

If the terms being defined refer to universals, then the hierarchy of universals from more general (animal, philosophy of science, furniture) to more specific (mammal, philosophy of biology, chair) should be reflected in the definitions of the terms that refer to these universals. Terms lower down in a taxonomic hierarchy should inherit from their parents all characteristics asserted to be true in the ontology. This ensures logical consistency in the definition of terms, clear demarcations amongst levels of abstractness within the ontology, and the possibility of automated reasoning.

*Figure 7: A Few Biomedical Relations*

**Foundational Relations**

(1) is_a meaning "is a subtype of" as in:
- DNA is_a nucleic acid
- photosynthesis is_a physiological process

(2) part_of meaning "is a part of" as in:
- nucleoplasm part_of nucleus
- neurotransmitter release part_of synaptic transmission

**Spatial Relations**

(3) located_in meaning "is located in" as in:
- intron located_in gene
- chlorophyll located_in thylakoid

(4) contained_in meaning "is contained in" as in:
- synaptic vesicle contained_in neuron
- cytosol contained_in cell compartment space

(5) adjacent_to meaning "is adjacent to" as in:
- Golgi apparatus adjacent_to endoplasmic reticulum
- periplasm adjacent_to plasma membrane

**Temporal Relations**

(6) transformation_of meaning "is a transformation of" as in:
- mature mRNA transformation_of pre-mRNA

- foetus transformation_of embryo

(7) derives_from meaning "derives from" as in:
- mammal derives_from gamete
- triple oxygen molecule derives_from oxygen molecule

(8) preceded_by meaning "is preceded by" as in:
- translation preceded_by transcription
- digestion preceded_by ingestion

**Participation Relations**

(9) has_participant meaning "has as a participant in its process" as in:
- death has_participant organism
- breathing has_participant thorax

(10) has_agent meaning "has as an agent in its process" as in:
- translation has_agent ribosome
- signal transduction *has_agent* receptor

Once the taxonomy is established using the *is_a* relations between terms, other relevant relations may be added to the domain ontology. For example, Figure 7 represents some significant relations one would find in many biomedical domain ontologies.

**Step 6: Regiment the Ontology to Ensure Logical Coherence and Accuracy of Information.**

The goal of regimentation is to develop a domain ontology that is logically coherent, unambiguous, and maximally correct when viewed in light of the current state of the relevant science, discipline, or area of study. This does not mean that the ontology must be complete. Rather, it should contain the terminological content of those parts of the relevant discipline which have become established as textbook knowledge. Coherent definitions are essential to constructing any domain ontology, and much of the effort involved in building domain ontologies consists in putting forward clearly defined terms. Whenever a definition of a term is proposed, a thorough attempt must be made to identify potential counter-examples.

Further, a defined term in the ontology should be *intersubstitutable* with its definition in such a way that the result is both (a) grammatically correct, and (b) truth-preserving. Thus, for example, in the *Foundational Model of Anatomy* (FMA) the extension of the term 'heart' should be identical with the collection of all those things which satisfy the definition: "organ with cavitated organ parts, which is continuous with the systemic and pulmonary arterial and venous trees" (see: http://sig.biostr. washington.edu/projects/fm/About FM.html). The proposition:

- The *heart* pumps blood

then means the same thing as:

- The *organ with cavitated organ parts, which is continuous with the systemic and pulmonary arterial and venous trees* pumps blood.

The intersubstitutability of a term, and its definition with regard to the truth-value of sentences in which they occur, is important for the human users of ontologies as well as for the computational systems that will be making inferences on their basis.

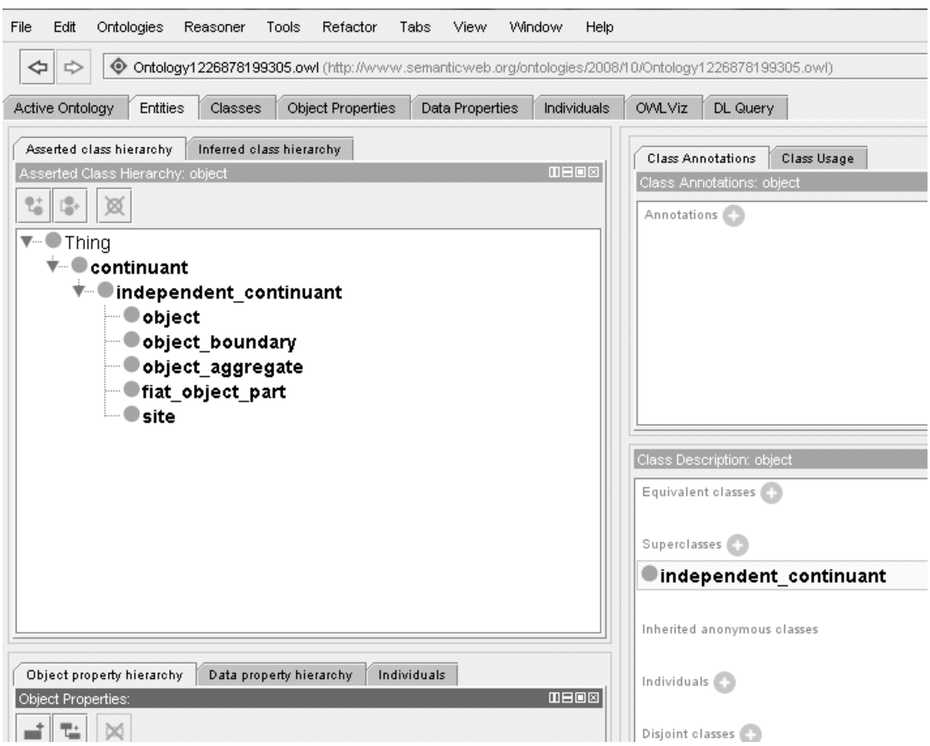
**Step 7: Seek Interoperability by Using a Common Formal Ontology.**

This step is crucial to de-siloing information and interoperability. Here, domain ontologists can link the terms in their domains to a formal ontology, like the *Standard Upper Merged Ontology* (SUMO), the *Descriptive Ontology for Linguistic and Cognitive Engineering* (DOLCE), or *Basic Formal Ontology* (BFO, see Figure 5).

**Step 8: Concretize the Ontology in a Computer Tractable Representational Artifact.**

This step is essentially a *syntactical* one, and involves translating the representations of universals and relations contained in the ontology into a form that is readable by a computer. Protégé is an open source ontology editor—available through http://protege.stanford.edu/—that is being used in the creation and editing of ontologies as computational artifacts. Figure 8 shows the results of importing BFO into Protégé (only *BFO:independent_continuant* entities shown), where it can serve as a starting point for further ontology content entry in consistent fashion as one moves further down the tree.

*Figure 8: A Screen Shot of Protégé*




**Step 9: Apply the Ontology, and Test the Results in a Computing Context.**

Finally there comes the process of applying the domain ontology to real data, as for example in the case of the *Gene Ontology Annotation* database (http://www.ebi.ac.uk/GOA).

## Developing a Philosopher's Résumé Domain Ontology

Philosophers, like anyone else, would like to have quick access to basic information about a person's work experience, publications, presentations, and the like, as would be found in a résumé or curriculum vita. In conversations about some philosopher's important work, it is typical to hear questions like: At what school is she now located? From where did she attain her doctorate? Who was her primary mentor? How many books and articles does she have published? As one can imagine, having such information about a philosopher readily available would be especially helpful for hiring committees and tenure review boards. We are now in a position to apply what has been spoken about so far in this paper by taking the first steps in developing a *philosopher's résumé domain ontology*. This application will be brief and skeletal in nature because of space limitations; however, it will be helpful for philosophers and others to see a concrete example of domain ontology-building at work.

*Figure 9: First Steps in Constructing a Domain Ontology*

*Step 1: Determine the Purpose of the Domain Ontology.*
*Step 2: Provide an Explicit Statement of the Intended Subject Matter.*
*Step 3: Determine the Most Basic Universals and Relations, Clearly and Coherently Defining Them.*
*Step 4: Use Aristotelian Structure When Formulating Definitions.*
*Step 5: Put the Universal Terms in a Taxonomic Hierarchy, Adding the Relevant Relations.*
*Step 6: Regiment the Ontology to Ensure Logical Coherence and Accuracy of Information.*
*Step 7: Seek Interoperability by Using a Formal Ontology such as Basic Formal Ontology.*
*Step 8: Concretize the Ontology in a Computer Tractable Representational Artifact.*
*Step 9: Apply the Ontology, and Test the Results in a Computing Context.*

In line with the first and second steps in building a domain ontology spoken about already (see Figure 9 for quick reference), this philosopher's résumé domain ontology serves primarily as an application ontology that will be used to mine and query information that one would find typically on a philosopher's résumé and/or curriculum vita. In line with the third and fourth steps, the following is a tentative list of only *some* of the universals and relations dealt with in this domain. Further, only a few of the definitions have been provided, as most of the definitions of the terms are intuitively obvious; however, some of the definitions may be controversial or debatable. Since this is a domain ontology that concerns, in part, individual philosophers, books, institutions, and the like, particular entities are included below the universals and relations. The example utilized below primarily has to do with Martin Heidegger's résumé, but one will readily see how it has general applicability to any philosopher, living or deceased.

### Universals:

person: =def. a human being that is conscious and the full bearer of rights and privileges in a society.

philosopher: =def. a role that a person has whereby that person is considered a practitioner of philosophy.

philosophy: =def. a discipline that studies matters concerning logic, metaphysics, epistemology, ethics, political philosophy, and their associated sub-disciplines.

philosophical discipline: =def...
logic: =def...
metaphysics:
epistemology:
ethics:
political philosophy:
phenomenology:
existentialism:
area of specialization:
doctorate in philosophy:
publication:
book:
edited book:
book chapter:
anthology:
article:
university:
philosophy department:
student:
teacher:
rank:
title:
postal address:
email address:
phone number:
fax number:
teaching fellowship:
postdoctoral fellowship:
conference presentation:
reviewer:
grant:
honor:
language proficiency:
German language proficiency:
French language proficiency:
Latin language proficiency:
Greek language proficiency:
English language proficiency:

### Relations Besides the *is_a* Subtype Relation:

*has_role:* =def. a relation between some object O and a role R, and the inverse of *role_of*, whereby the object exercises some optional activity in a special natural, social, or institutional set of circumstances.

*has_area_of_specialization:* =def. a relation between a person P and an area of specialization A, and the inverse of *area of specialization_of*, whereby the person can claim expert knowledge concerning an area of study as demonstrated by successful publications, presentations, and teaching.

*has_birthdate:* =def...
*has_deathdate:* =def...

*has_student:*
*has_language_proficiency:*
*employed_at:*
*has_postal_address:*
*has_email_address:*
*has_phone_number:*
*has_fax_number:*
*colleague_of:*
*received_PhD_in_Philosophy_from:*
*has_area_of_concentration:*
*has_rank:*
*has_title:*
*has_postdoctoral_fellowship_at:*
*has_teaching_fellowship_at:*
*has_book:*
*has_edited_book:*
*has_book_chapter:*
*has_article:*
*has_conference_presentation:*
*has_invited_speaker_presentation:*
*has_award:*
*has_honor:*
*has_grant:*
*reviewer_for:*

**Particular Entities:**

Martin Heidegger: =def...
Edmund Husserl: =def...
Hans-Georg Gadamer:
Hannah Arendt:
Rudolf Bultmann:
*Cartesian Meditations:*
*Being and Time:*
*Truth and Method:*
*The Human Condition:*
University of Freiburg:
University of Marburg:
University of Heidelberg:

In line with the fifth and sixth steps, we can now put some of the above terms in taxonomic hierarchies with the relevant relations, as well as ensure logical coherency and accuracy of information. Figure 10 shows a graphing of a few of the entities and relations as concerns Martin Heidegger.

***is_a* Relation:**

philosopher *is_a* role
student *is_a* role
teacher *is_a* role
philosophy *is_a* discipline
book *is_a* publication
article *is_a* publication
edited book *is_a* publication
book chapter *is_a* publication
anthology *is_a* publication
English language proficiency *is_a* language proficiency
German language proficiency *is_a* language proficiency
French language proficiency *is_a* language proficiency
Latin language proficiency *is_a* language proficiency
Greek language proficiency *is_a* language proficiency

logic *is_a* philosophical discipline
metaphysics *is_a* philosophical discipline
epistemology *is_a* philosophical discipline
ethics *is_a* philosophical discipline
political philosophy *is_a* philosophical discipline
phenomenology *is_a* philosophical discipline
existentialism *is_a* philosophical discipline

***has_role*** **Relation:**
Edmund Husserl *has_role* philosopher
Martin Heidegger *has_role* philosopher
Hans-Georg Gadamer *has_role* philosopher
Hannah Arendt *has_role* philosopher

***has_area_of_specialization*** **Relation:**
Edmund Husserl *has_area_of_specialization* phenomenology
Martin Heideger *has_area_of_specialization* phenomenology
Martin Heideger *has_area_of_specialization* existentialism
Hans-Georg Gadamer *has_area_of_specialization* existentialism
Hannah Arendt *has_area_of_specialization* existentialism

***has_birthdate*** **Relation:**
Edmund Husserl *has_birthdate* April 8, 1859
Martin Heidegger *has_birthdate* September 26, 1889
Hans-Georg Gadamer *has_birthdate* February 11, 1900
Hannah Arendt *has_birthdate* October 14, 1906

***has_deathdate*** **Relation:**
Edmund Husserl *has_deathdate* April 26, 1938
Martin Heidegger *has_deathdate* May 26, 1976
Hans-Georg Gadamer *has_deathdate* March 13, 2002
Hannah Arendt *has_deathdate* December 4, 1975

**has_student Relation:**
Edmund Husserl *has_student* Martin Heidegger
Martin Heidegger *has_student* Hans-Georg Gadamer
Martin Heidegger *has_student* Hannah Arendt

**has_language_proficiency Relation:**
Edmund Husserl *has_language_proficiency* German language proficiency
Martin Heidegger *has_language_proficiency* German language proficiency
Hans-Georg Gadamer *has_language_proficiency* German language proficiency
Hannah Arendt *has_language_proficiency* German language proficiency

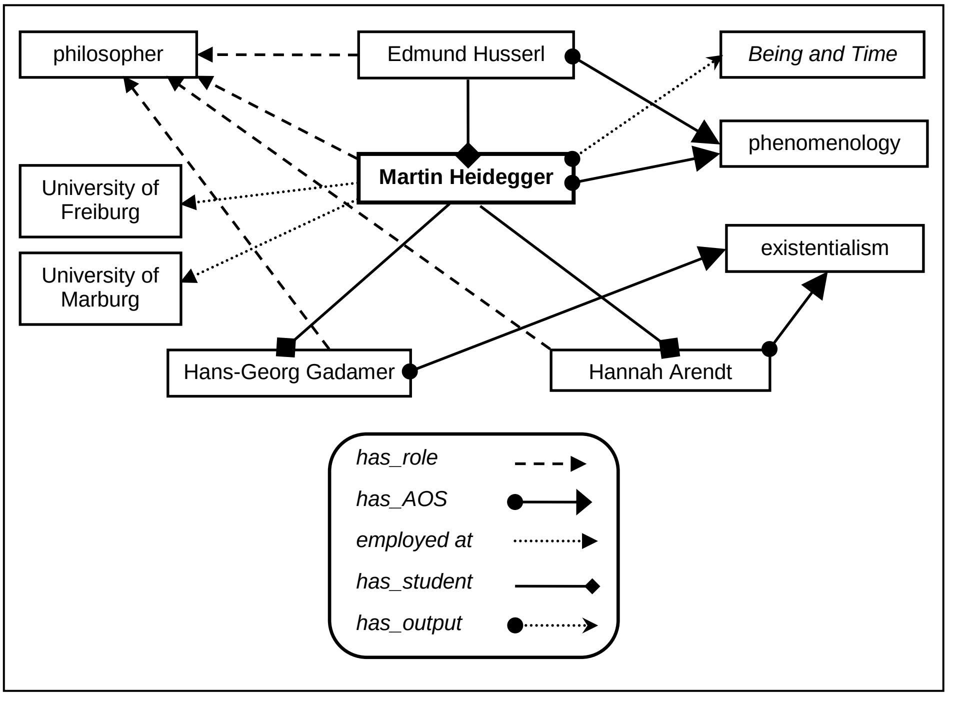
***has_book*** **Relation:**
Edmund Husserl *has_book* Cartesian Meditations
Martin Heidegger *has_book* Being and Time
Hans-Georg Gadamer *has_book* Truth and Method
Hannah Arendt *has_book* The Human Condition

***employed_at*** **Relation:**
Martin Heidegger *employed_at* University of Freiburg
Martin Heidegger *employed_at* University of Marburg

***colleague_of*** **Relation:**
Martin Heidegger *colleague_of* Rudolf Bultmann

***received_PhD_in_Philosophy_from*** **Relation:**
Martin Heidegger *received_PhD _in_Philosophy_from* University of Freiburg
Hannah Arendt *received_PhD_in_Philosophy_from* University of Heidelberg
Hans-Georg Gadamer *received_PhD_in_Philosophy_from* University of Marburg

*Figure 10: Part of a Philosopher's Résumé Domain Ontology*



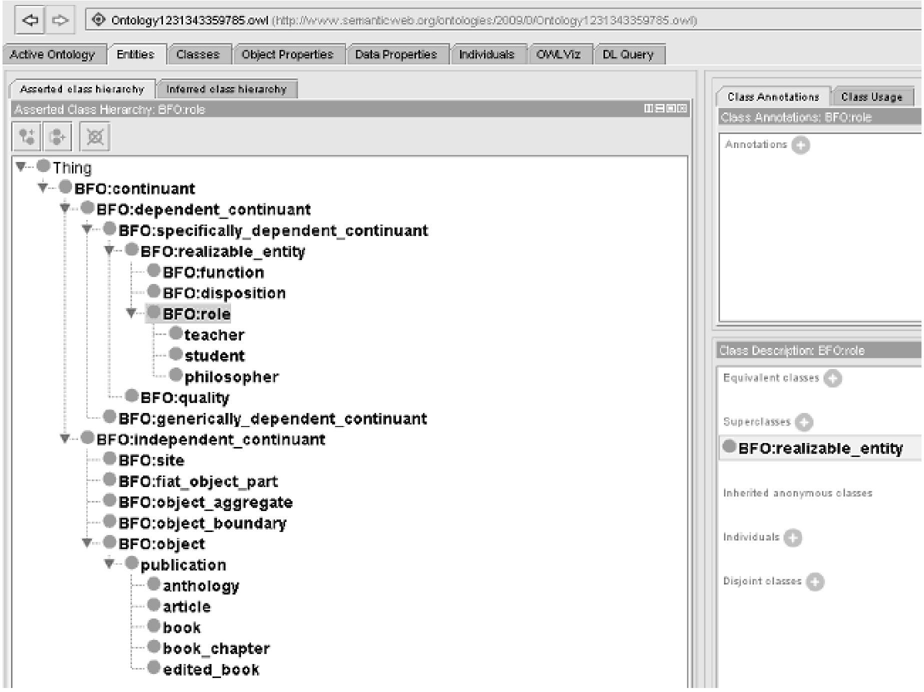

In line with the seventh step of seeking interoperability, not only can we link these terms in the philosopher's résumé domain ontology with an upper level ontology like Basic Formal Ontology (see Figure 11), but we can also create links between terms in this domain and other databases of information relevant to the philosophical disciplines. For example, we can link Edmund Husserl or phenomenology to the *Stanford Encyclopedia of Philosophy*, *Internet Encyclopedia of Philosophy*, or *A Taxonomy of Philosophy*.

In line with the eighth and ninth steps, we can use Protégé to classify the entities and relationships in our philosopher's résumé domain ontology and make it available on the World Wide Web for anyone to utilize and scrutinize (see Figure 11).

## Conclusion

Ontology is not just for philosophers anymore, as information-based science has spawned myriad domain ontologies and formal ontologies which are being put to use in a host of different ways. Yet, philosophy and philosophical methods are now assisting informaticians in realizing the goals of information accessibility and shareability. And while few philosophers are as yet engaging in this new applied ontology, it presents an increasingly important area where philosophical ideas can be applied and bring real benefits to the development of various disciplines, including philosophy itself.

*Figure 11: Screen Shot of Protégé Showing Part of a Philosopher's Résumé Domain Ontology*




## References


Arp, R. (2008). Realism and antirealism in informatics ontologies. *The American Philosophical Association: Philosophy and Computers*, 8*(*2).

——— (2009). Philosophical ontology. Domain ontology. Formal ontology. In J. Williamson & F. Russo (eds.), *Key Terms in Logic* (London: Continuum Press), pp.122–123.

Arp, R., & Smith, B. (2008). Ontologies of cellular networks. *Science Signaling*, *1*, mr2.

Arp, R., Chhem, R., Romagnoli, C., & Overton, J. (2008). Radiological and biomedical knowledge integration: The ontological way. In R. Chhem, K. Hibbert, and T. Van Deven (eds.), *Scholarship in radiology education* (Berlin: Springer), pp.87–104.

Ceusters, W., Capolupo, M., De Moor, G., & Devlies, J. (2008). Introducing realist ontology for the representation of adverse events. In C. Eschenbach & M. Gruninger (eds.), *Formal ontology in information systems* (Amsterdam: IOS Press). Available from: http://www.org.buffalo.edu/RTU/papers/Ceusters FOIS20081NEW-VERSION.pdf.

Ceusters, W., Smith, B., Kumar, A., & Dhaen, C. (2004). Mistakes in medical ontologies: Where do they come from and how can they be detected? *Studies in Health Technology and Informatics*, 102: 145–164.

Domingue, J., & Motta, E. (1999). A knowledge-based news server supporting ontology-driven story enrichment and knowledge retrieval. In D. Fensel & R. Studer (eds.), *Knowledge Acquisition, modeling and management* (Berlin: Springer), pp.104-112.

Feigenbaum, L., Herman, I., Hongsermeier, T., Neumann, E., & Stephens, S. (2007). The Semantic Web in action. *Scientific American*, 297: 90–97.

Grenon, P., & Smith, B. (2004). SNAP and SPAN: Towards dynamic spatial ontology. *Spatial Cognition and Computation*, 1: 1–10.

Gruber, T. (1993). A translation approach to portable ontologies. *Knowledge Acquisition*, 5, 199–220.

Husserl, E. (1900). *Logische Untersuchungen. Erste Teil: Prolegomena zur reinen Logik*. Tübingen: Max Niemeyer.

——— (1901). *Logische Untersuchungen. Zweite Teil: Untersuchungen zur Phänomenologie und Theorie der Erkenntnis*. Tübingen: Max Niemeyer.

Kim, J., Choi, B., Shin, H., & Kim, H. (2007). A methodology for constructing of philosophy ontology based on philosophical texts. *Computer Standards & Interfaces*, 29: 302–315.

Lowe, E. (2006). *The four category ontology: A metaphysical foundation for natural science*. Oxford: Oxford University Press.

Luenberger, D. (2006). *Information Science*. Princeton: Princeton University Press.

Luger, G. (2008). *Artificial Intelligence: Structures and strategies for complex problem solving*. New York: Addison-Wesley.

Polanski, A., & Kimmel, M. (2007). *Bioinformatics*. Berlin: Springer.

Rosse, C., & Mejino, J. (2003). A reference ontology for biomedical informatics: the Foundational Model of Anatomy. *Journal of Biomedical Informatics*, 36: 478–500.

Smith, B. (2003). Ontology. In L. Floridi (ed.), *Blackwell Guide to the Philosophy of Computing and Information*, pp.155–166. Malden, MA: Blackwell.

Smith, B., & Smith, D. (eds.) (1995). *The Cambridge Companion to Husserl*. Cambridge: Cambridge University Press.

Smith, B., Kusnierczyk, W., Schober, D., & Ceusters, W. (2006). Towards a reference terminology for ontology research and development in the biomedical domain. *Proceedings of KR-MED*, 1(7).

Smith, B., Ashburner, M., Rosse, C., Bard, J., Bug, W., Ceusters, W., *et al.*, (2007). The OBO Foundry: Coordinated evolution of ontologies to support biomedical data integration. *Nature Biotechnology*, 25, 1251–1255.

Smith, B., Ceusters, W., Klagges, B., Köhler, J., Kumar, A., Lomax, J. (2005). Relations in biomedical ontologies. *Genome Biology*, 6: R46.

Spear, A. (2006). *Ontology for the Twenty First Century: An introduction with recommendations*. Available from: http://www.ifomis.org/bfo/manual.pdf

Vizenor, L., Smith, B., & Ceusters, W. (2004). Foundation for the electronic health record: An ontological analysis of the HL7's reference information model. Available from: http://ontology. buffalo.edu/HL7/index.htm.

## Acknowledgements

I thank Barry Smith, Werner Ceusters, Andrew Spear, Lowell Vizenor, Colin Allen, and Anthony Beavers for material, comments, and fruitful conversations concerning this paper. This work was funded by the *National Institutes of Health* through the NIH Roadmap for Medical Research, Grant 1 U 54 HG004028. Information on the *National Centers for Biomedical Computing* can be found at: http://nihroadmap.nih.gov/bioinformatics

## About the Author

**Robert Arp** has a Ph.D in Philosophy from *Saint Louis University*, and is now working as a contractor building ontologies for the US Air Force. He also has interests in philosophy and popular culture, and philosophy of biology. His publications include *Scenario Visualization: An Evolutionary Account of Creative Problem Solving* (MIT Press, 2008); *Philosophy of Biology: An Anthology* (Wiley-Blackwell, 2009), as co-editor with Alex Rosenberg; *Contemporary Debates in Philosophy of Biology* (Wiley-Blackwell, 2009), as co-editor with Francisco Ayala.

robertarp320@gmail.com