4-2016

A black art: Ontology, data, and the Tower of Babel problem

Andrew J. Iliadis
Purdue University

Graduate School Form
30 Updated 12/26/2015

PURDUE UNIVERSITY
GRADUATE SCHOOL
Thesis/Dissertation Acceptance

This is to certify that the thesis/dissertation prepared

By Andrew Iliadis

Entitled
A Black Art: Ontology, Data, and the Tower of Babel Problem

For the degree of Doctor of Philosophy

Is approved by the final examining committee:

Ashley R. Kelly
Co-chair

Daniel W. Smith
Co-chair

Sorin A. Matei

Daniel R. Kelly

To the best of my knowledge and as understood by the student in the Thesis/Dissertation Agreement, Publication Delay, and Certification Disclaimer (Graduate School Form 32), this thesis/dissertation adheres to the provisions of Purdue University's "Policy of Integrity in Research" and the use of copyright material.

Approved by Major Professor(s): Ashley R. Kelly and Daniel W. Smith

Approved by: Melanie Morgan 4/16/2016

Head of the Departmental Graduate Program Date

A BLACK ART: ONTOLOGY, DATA, AND THE TOWER OF BABEL PROBLEM

A Dissertation

Submitted to the Faculty

of

Purdue University

by

Andrew J Iliadis

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

May 2016

Purdue University

West Lafayette, Indiana

For my parents

ACKNOWLEDGEMENTS

I would like to acknowledge the support of the Brian Lamb School of Communication and the Department of Philosophy at Purdue University. I would also like to thank my co-advisors and committee members, including Dr. Ashley R. Kelly, Dr. Daniel W. Smith, Dr. Sorin A. Matei, and Dr. Daniel R. Kelly. A special thank you to my parents, Cathy and Jim, and to my brother, Michael. I could not have done this without your support.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

Figure Page

LIST OF ABBREVIATIONS

Buffalo Center for Ontological Research (BCOR)

Basic Formal Ontology (BFO)

New York State Center of Excellence in Bioinformatics & Life Sciences (CBLS)

Chairman of the Joint Chiefs of Staff Instruction (CJCSI)

Descriptive Ontology for Linguistic and Cognitive Engineering (DOLCE)

Information Artifact Ontology (IAO)

International Association for Ontology and Its Applications (IAOA)

Institute for Formal Ontology and Medical Information Science (IFOMIS)

Informational Structural Realism (ISR)

Laboratory for Applied Ontology (LAO)

National Center for Biomedical Ontology (NCBO)

National Center for Ontological Research (NCOR)

Open Biological and Biomedical Ontologies Foundry (OBO)

Ontology Research Group (ORG)

Web Ontology Language (OWL)

Tower of Babel Problem (ToB)

ABSTRACT

Iliadis, Andrew J. Ph.D., Purdue University, May 2016. A Black Art: Ontology, Data, and the Tower of Babel Problem. Major Professors: Ashley R. Kelly and Daniel W. Smith.

Computational ontologies are a new type of emerging scientific media (Smith, 2016) that process large quantities of heterogeneous data about portions of reality. Applied computational ontologies are used for semantically integrating (Heiler, 1995; Pileggi & Fernandez-Llatas, 2012) divergent data to represent reality and in so doing applied computational ontologies alter conceptions of materiality and produce new realities based on levels of informational granularity and abstraction (Floridi, 2011), resulting in a new type of informational ontology (Iliadis, 2013) the critical analysis of which requires new methods and frameworks. Currently, there is a lack of literature addressing the theoretical, social, and critical dimensions of such informational ontologies, applied computational ontologies, and the interdisciplinary communities of practice (Brown & Duguid, 1991; Wenger, 1998) that produce them. This dissertation fills a lacuna in communicative work in an emerging subfield of Science and Technology Studies (Latour & Woolgar, 1979) known as Critical Data Studies (boyd & Crawford, 2012; Dalton & Thatcher, 2014; Kitchin & Lauriault, 2014) by adopting a critical framework to analyze the systems of thought that inform applied computational ontology while offering insight into its realism-based methods and philosophical frameworks to gauge their ethical

import. Since the early 1990s, computational ontologies have been used to organize massive amounts of heterogeneous data by individuating reality into computable parts, attributes, and relations. This dissertation provides a theory of computational ontologies as technologies of individuation (Simondon, 2005) that translate disparate data to produce informational cohesion. By technologies of individuation I mean engineered artifacts whose purpose is to partition portions of reality into computable informational objects. I argue that data are metastable entities and that computational ontologies restrain heterogeneous data via a process of translation to produce semantic interoperability. In this way, I show that computational ontologies effectively re-ontologize (Floridi, 2013) and produce reality and thus that have ethical consequences, specifically in terms of their application to social reality and social ontology (Searle, 2006). I use the Basic Formal Ontology (Arp, Smith, & Spear, 2015) – the world's most widely used upper-level ontology – as a case study and analyze its methods and ensuing ethical issues concerning its social application in the Military Ontology before recommending an ethical framework. "Ontology" is a term that is used in philosophy and computer science in related but different ways—philosophical ontology typically concerns metaphysics while computational ontology typically concerns databases. This dissertation provides a critical history and theory of ontology and the interdisciplinary teams of researchers that came to adopt methods from philosophical ontology to build, persuade, and reason with applied computational ontology. Following a critical communication approach, I define applied computational ontology construction as a solution to a communication problem among scientists who seek to create semantic interoperability among data and argue that applied ontology is philosophical, informational in nature, and communicatively constituted

(McPhee & Zaug, 2000). The primary aim is to explain how philosophy informs applied computational ontology while showing how such ontologies became instantiated in material organizations, how to study them, and describe their ethical implications.

CHAPTER 1. THE TOWER OF BABEL PROBLEM

1.1 Introduction

Applied computational ontologies are quickly becoming one of the most complex and pervasive forms of emerging scientific media (Smith, 2016) in the world and stand to revolutionize entire industries and domains of social life (Staab & Studer, 2009). Invisible to the practices of everyday life (de Certeau, 1984), applied computational ontologies process the entities, attributes, and relations of portions of reality in various social, governmental, economic, and scientific contexts (Nissenbaum, 2011) into computable objects separated by invisible granular partitions, producing a new way of understanding materiality that can be called informational ontology (Iliadis, 2013). Such applied computational ontologies exert forms of power and control that affect everyday life yet applied computational ontologies are rarely studied outside of the domains of computer and information science and should be exposed to qualitative and ethical analysis stemming from work in fields such as Philosophy, Communication, Media Studies, and Science and Technology Studies.

Critical Data Studies (CDS) is a new subfield of Science and Technology Studies (boyd & Crawford, 2012; Dalton & Thatcher, 2014; Kitchin & Lauriault, 2014), that focuses on the unique historical, infrastructural, methodological, epistemological, and ethical frameworks involved in understanding data, their structures, and data's ensuing impact on social life. This dissertation takes a CDS approach to the study of applied computational ontologies and the informational ontologies they produce by showing how a specific branch of philosophical ontology known as ontological realism directly informs a popular applied computational ontology called the Basic Formal Ontology (BFO). I explain the logic and methods that contributed to the BFO's success in becoming the most widely-used upper-level ontology in the world. Through long form unstructured ethnographic interview data gathered from six interviews with key individuals involved in the creation of the BFO at the National Center for Ontological Research, I explain the BFO's rise in the domain of natural science and then attend to the BFO's ethical implications in the second half of the dissertation in the context of social ontology. Proponents of the BFO present the BFO as a purely realism-based and scientific ontology. I show that the BFO is applied to social contexts when it is used by entities such as the military and thus that the BFO is used for intelligence sharing and social organization and control.

The dissertation has two parts and five chapters. The first part provides an overview of the relevant systems of thought and the historical and theoretical connections between the philosophical and scientific literature on realism-based approaches to ontology while the second part offers a modern empirical case study of the BFO and its application to the Military Ontology. I begin the dissertation by providing a theory of informational ontology and technologies of individuation and introduce a communication problem regarding semantic interoperability in computational ontology known as the Tower of Babel problem (Smith, 2004a; Blass, Gurevich, & Hudis, 2007) (chapter one). The first three chapters cover a broad spectrum of philosophical theories about ontology and provides evidence of the value of one specific branch of ontology over another (the BFO's realism-based ontology over concept-based ontology) (chapter two) before recommending a communicative approach to ontology studies by examining some of the main groups, institutions, individuals, and communities of practice involved in the BFO's applied ontology work (chapter three). The second part offers a case study of science teams involved in the construction, use, and maintenance of the BFO (chapter four) and introduces the field of CDS to discuss the BFO's potential ethical issues in applied social ontologies such as the Military Ontology (chapter five). The BFO-based Military Ontology project was a direct response to a Chairman of the Joint Chiefs of Staff Instruction (CJCSI) entitled "Horizontal Integration (HI) of Warfighter Intelligence." I show how the CJCSI's mandate to construct the Military Ontology is an example of BFO's application to social and governmental spheres, suggesting that the realism-based methods of the BFO should be reviewed when applied in the context of social ontology.

Borrowing from the French philosopher Gilbert Simondon's theory of individuation (2005), I describe applied computational ontologies such as the BFO as technologies of individuation that translate heterogeneous data to produce semantic interoperability, thus contributing to a reconceptualization of materiality as consisting of levels of informational abstraction. As technologies of individuation, applied computational ontologies not only re-ontologize (Floridi, 2013) reality but also produce new realities that have ethical consequences for social groups and individuals. Computational ontologies are tools that shape reality according to a new understanding of informational ontology and thus should be reviewed to measure their ethical impact.

A Definition of Ontology. Following from Husserl (1900/2001; 1901/2001), Ingarden (1947), Quine (1948), and Simondon (2005), I define ‘ontology’ as the philosophical and comprehensive study of the objects, entities, attributes, events, processes, relations, and structures that exist in reality and the principles through which they can be revealed. While there are different approaches to practicing ontology, the branch discussed here is tied to realism.

A Definition of Informational Ontology. In Iliadis and Russo (forthcoming), we build on Smith (2004) and Floridi’s (2008a) notion of informational ontology, which we define as the multiple levels of informational granularity and abstraction though which reality is accessed and revealed. Informational ontology means that the ultimate nature of reality is best understood through informational structural realism (Floridi, 2011).

General Questions. Four general questions that orient this dissertation are: Has ontology influenced scientific knowledge production? How have ontologies changed scientific practice? How has ontology changed scientific communication? Are ontologies a new way to communicate?

Situated Questions. Floridi (2008a) suggests that “the ultimate nature of reality is informational” and can be understood through ISR. Smith and Ceusters (2010) suggest an ontological realism approach to applied computational ontologies as a way to process that informational reality. Russo and I (forthcoming) take up these issues to critically theorize informational ontology and computational ontologies using a CDS framework.

Research Questions. The five research questions that inform this study are: RQ #1 Is there a connection between philosophical ontology and computational ontology? RQ #2 Does Smith’s claim that ontological realism is the best way to practice ontology hold up? RQ #3 Are there new types of ontologies that adopt the method of ontological realism? RQ #4 If there are, how are they applied and are there any ethical consequences? RQ #5 What might this tell us about ontology? RQ #1 is answered in chapters one and two, RQ #2 is answered in chapters three and four, RQ #3 is answered in chapter four, and RQ #4 is answered in chapter five. RQ #5 is a general research question based on the previous four and is answered throughout the dissertation.

Theory-building and Methods. I use ontological theories developed by philosophers such as Aristotle, Husserl, Ingarden, Quine, Simondon, Smith, and Floridi to identify key moves in the history of ontology and to provide a critical analysis of ontology's systems of thought, communities of practice, and practical application. I take applied computational ontologies as artifacts to help identify these changes, through the lens of CDS, with interviews and ethnography supplementing my analysis. Case study research is used to help uncover the complexity of these applied ontological issues.

The main take home message of the dissertation is that there are ethical implications to applied computational ontologies for social organization and control.

1.2 Philosophical and Computational Ontology

"Ontology" is a compound word that combines the Greek *ontos* (being, that which is) with *logia* (science, study, or theory). The first recorded known usage is from the Latin *ontologia* and appears in the German scholastic philosopher Rudolph Goclenius' (1547-1628) *Lexicon Philosophicum* of 1613. Medieval scholastics like Thomas Aquinas (1225-1274) and Duns Scotus (1266-1308) developed works in religious philosophical ontology based on Aristotelian metaphysics to describe entities, attributes, and relations that exist. The English philologist and lexicographer Nathan Bailey famously described ontology as "an Account of being in the Abstract" in *An Universal Etymological English Dictionary* (1721). Similarly, in *A Fragment on Ontology* (1843), Jeremy Bentham writes that the "field of ontology, or as it may otherwise be termed, the field of supremely abstract entities, is a yet untrodden labyrinth."

From a broad, philosophical perspective, 'ontology' (in philosophy 'ontology' is often treated as a singular noun whereas in computer and information science it is more often plural) is the study of what exists. 'Ontology' is sometimes conflated with 'metaphysics,' a word often erroneously attributed to Aristotle (384-322 BC)—'meta' comes from the Greek word for 'after' and in Aristotle's case this was meant to refer to his works that were anthologized after his works on physics. Aristotle referred to his work on the nature of 'being *qua* being' as a type of 'first philosophy.' Ontology remains an important area of philosophy and has splintered into many permutations and subgroups divided along various longstanding dichotomies (philosophical categories represented by such terms as universal and particular, substance and accident, abstract and concrete, essence and existence, and so forth).

In his famous 1948 essay "On What There Is," the American philosopher and logician Willard Van Orman Quine (1908-2000) wrote: "Our acceptance of an ontology is, I think, similar in principle to our acceptance of a scientific theory" (p. 35). On Quine's theory, philosophical ontologies represent our closest representation for how 'things' are 'out there' in the world by committing to the most accurate account of the objects, entities, events, processes, relations, and structures that exist in reality as identified by science. This version of philosophical ontology can be described as tied to the scientific method insofar as it seeks not an explanation but rather a description of reality in terms of a classification of entities that is exhaustive, however Quine's is by no means the only account. The relationship between philosophical ontology and computational ontology has received little attention by both fields and this is unfortunate for numerous reasons, not least of which is that philosophy contains the necessary conceptual tools for successful ontology engineering.

Computational ontology is a burgeoning field that stands to revolutionize data organization at the intersection of a variety of disciplines, from bioinformatics (Bodenreider & Stevens, 2006) to management (Allen & March, 2012), military intelligence (Dragos, 2013) to farming (Sivamani, Bae, & Cho, 2013). In this chapter, I provide a theory of informational ontology (reality understood as consisting of levels of informational abstraction) and technologies of individuation (computational ontologies used to process that reality) based on the work of the French philosopher Gilbert Simondon (1924-1989). I show how computational ontology is historically linked to philosophical ontology and discuss a difficult problem in modern applied ontology known as the Tower of Babel problem (ToB) which concerns the difficulty of integrating heterogeneous datasets to produce semantic interoperability among data. I introduce a brand of realism-based applied ontology developed by leading ontology experts Dr. Barry Smith and Dr. Werner Ceusters (University at Buffalo) as a solution to ToB and defend it from longstanding concept-based approaches to applied ontology inherited from early scientific research in artificial intelligence (AI). Excerpts from long form unstructured interviews with ontologists is presented throughout as evidence of the ToB problem and potential of the realism-based applied ontology framework.

Ontologies are an attempt at solving ToB via the creation of formal taxonomies for use among disparate domains of application. There has been a widely-recognized need for “practical methodologies and technologies, which can assist a variety of user types with ontology development” (Suárez-Figueroa, Gómez-Pérez, Motta, & Gangemi, 2012a; 2012b, p. 1). Multiple international organizations and centers have been created to assist in ontology engineering and the development of methods for processing disparate domains of data, including the European Centre for Ontological Research, the USA’s National Center for Ontological Research, and the Buffalo Center for Ontological Research. There are journals, societies, and conferences dedicated to ontology, including *Applied Ontology: An Interdisciplinary Journal of Ontological Analysis and Conceptual Modeling*, the International Association for Ontology and its Applications, and the International Conference on Formal Ontology in Information Systems, among others.

Philosophical ontology asks general questions that become practically applied in computational ontology. Philosophical ontology questions how universals might be related to particulars and how events can exist which bring multiple distinct entities together (for example, do such events produce new ‘super’ entities?). It asks questions that attempt to reach further than this into the metaphysics of topics such as God and morality. Putting such questions aside, ontology asks questions related to how materiality is made up and the relations under which materiality may be constrained. It asks the most general and basic questions that can be asked about materiality in an attempt to, as Plato put it, ‘carve nature at its joins’ with the goal of offering a plain yet comprehensive thesis for how the world is organized. Generally speaking, there are two parts to any ontological project; the first concerns what entities exist and of what attributes those entities are comprised while the second is concerned with the general relations that inhere to those entities. To complicate matters further, ontology must also give an account of things like holes, nothingness, or absences. Recalling the final line of Wallace Stevens' poem "The Snowman" (1921): ontology must address "the nothing that is not there and the nothing that is." For example, how should *aortic dissection* be categorized in an ontology?

One of the main problems in starting an inquest into philosophical ontology concerns the meta-ontological problem of ontological commitment. To make a decision on what might be said to exist (or not exist) philosophers must first make some type of baseline ontological commitment which allows them to build their ontology from the ground up. An ontology must start somewhere and make a commitment to an entity or relation if it is to grow; to make an ontological commitment is to plant the first stake in the ground that will dictate what can be included in the ontology from there on out. Various types of ontological commitments exist among philosophers and some of them offer irreconcilable claims. Following Chalmers (2009), realist philosophers might subscribe to the commonsense notion that entities such as chairs and bridges exist (this can be called *folk ontology*) while antirealists might say that such entities strictly speaking do not exist and that only, say, our concepts of them do.

In the same classic essay, Quine wrote that a theory "is committed to those and only those entities to which the bound variables of the theory must be capable of referring in order that the affirmations made in the theory be true" (1948, p. 33). Quine's highlighting of bound variables in ontological commitment is important to debates among philosophers and computer and information scientists in that it forces ontologists to make coherent decisions about what should be allowed in any given ontology using first-order logic (first-order logic uses quantified variables in place of objects). The underlying assumption is that at bottom there must be some kind of baseline commitment to an entity that makes the proposed ontology's internal logic uniformly veridical when variables refer to that entity. The bound ontological commitment endorsed by Quine has been adopted in some types of ontology engineering in the natural sciences (Oderberg, 2013) as it has enabled researchers to hypothesize and test theories related to naturally occurring entities and relations. Such researchers create concrete structures (computational ontologies) when attempting to model the world and in doing so they create what some have called a realist ontology (Weisberg, 2013) that contributes to scientific discovery.

Realism-based (Smith & Ceusters, 2010) and informational (Floridi, 2011) ontology can be referred to as informational ontology (Iliadis, 2013). Informational ontology borrows from the work of philosophical ontologists such as Husserl (1900/2001; 1901/2001), Ingarden (1947/2013), Quine (1960), and Smith (2004a)—philosophies that pay careful attention to various levels of informational abstraction according to the rules of first-order logic. Briefly, first-order logic is a type of reasoning involving symbols where every expression can be distilled to a subject and a predicate that modifies or defines the properties of the subject. Predicates in first-order logic must refer to a single subject and as such involve quantified variables over possibly non-quantifiable, non-logical objects and typically do not include universals (the ontological realism endorsed by ontologists such as Smith does). Such informational ontologies from philosophy are used by scientists today to conduct research in the natural sciences.

1.3 Ontology as Communication

After a philosophical ontology has been selected and before that ontological framework is expressed in an ontology language, researches argue and reason amongst themselves as they refine the ontology. This communicative middle ground is often overlooked in ontology research and by communication scholars or philosophers interested in the history of computer science; there is plenty of literature on the history of philosophical ontology and many volumes exist on practical computational ontologies. Apart from these, there is a noticeable lack of literature on what happens in-between these two realms in the area where actual researchers must settle on an ontological commitment and begin to reason and persuade each other through appeals to different ontological frameworks so that they may eventually agree on a standard approach that will be implemented. This is curious given the degree to which engineers and computer scientists often frame ontology problems in terms of communication problems. Communication problems and issues arise early in ontology-building, which should be treated as a communicative and open process that should be accessible to a variety of users.

Ontologies provide shared vocabularies that support communication among agents (Jakus, Milutinović, Omerović, & Tomažič, 2013) and as knowledge representations ontologies provide a medium for the human expression and communication of datasets and infrastructures (Davis, Shrobe, & Szolovits, 1993). Ontologies have rarely been studied as mediums of human data communication – that is, not only as constructs for AI and machine learning but also as mediums for human communication – yet they are a means to enable one specialized knowledge group to interact with another; human communication is central to the ontological enterprise in a way that reaches beyond technical specifications and abstract philosophical theorizing. Users must be able to understand the meanings produced by other humans in ontology-building and in this ontology is partly engaged in the practice of human communication.

As communicatively constituted organizations (McPhee & Zaug, 2000; Putnam, Nicotera, & McPhee, 2009; Putnam & Nicotera, 2009; 2010), ontologies are the product of communal deliberation, reasoning, and decision-making. Ontologies are co-constituted and co-constructed and must be communicatively maintained as their upkeep depends on groups of different editors and users. Bowker, Baker, Millerand, and Ribes (2010) highlight the role of such editors and users alongside technological and philosophical specifications since "who's in charge of crafting the ontology remains at stake" (p. 102). The notion of community is central to understanding ontologies as constituted artifacts; ontology work involves "taking knowledge out of a closed community of practice and allowing for its reuse and reshaping by others in different fields" (p. 109). Ontology engineering is combined with background work in identifying a larger community of future users (p.110). On this point, there have been calls for greater collaboration between philosophical ontologists and computer and information scientists (Smith & Welty, 2001). As a quintessential team science (Bozeman & Boardman, 2014), communication research in ontology construction should depend on talking to researchers who are actually involved in building ontologies. The life of ontology building involves ontologists who engage in daily decision-making and reasoning practices. To sufficiently understand ontologies as a form of communication entails looking beyond their technological specifications to the teams of researchers who are involved in their construction and maintenance; following early works in science and technology studies (Latour & Woolgar, 1979), attention should be placed on the interactions among ontologists and their communications in addition to technological and philosophical considerations. Ontology development is predominantly manual and laborious—though recent research in ontology expansion has attempted to automate this task (Pesquita & Couto, 2012). Full ontology automation will likely never be possible (partial ontology automation is a possibility) and ontologists continue to debate even minor rules and regulations in their ontologies, thus placing communication as central to the ontology-building process.

In short, ontology and communication are linked in at least three ways; ontologies are communicatively constituted, they facilitate and allow for data communication, and they change the way communication scholarship is practiced. Part of this dissertation is dedicated to exploring the relationship between ontology and communication, with the final aim of opening ontological practice up to fields outside of philosophy and science, including media, technology, and society.

Craig (1999) outlines a variety of communication domaines, including the rhetorical, semiotic, phenomenological, cybernetic, sociopsychological, sociocultural, and critical. To this one might add the ontological, meaning the communicative way that scholars and researchers effect change in materiality by way of applied ontology practice and maintenance. Ontological commitment may affect issues in communication scholarship; take for example the notion that "commitment to a particular ontology will influence one's epistemology and the attendant research methodology and protocol" (Arneson, 2009, p. 696).

1.4 Ontologies of Life and Death

Imagine the following scenario.

Bob has a family doctor named Alice. As a general practitioner, Alice examines Bob during his regular checkups and provides Bob with prescriptions when he is sick. Alice knows that Bob has been feeling nervous lately (for reasons that Bob cannot quite explain). After a heavy dinner and three glasses of wine, Bob begins to stammer and experiences a pronounced fluttering in his chest while finding it difficult to breathe. His wife rushes him to the emergency room. The doctors on call administer beta blockers, diagnose Bob with tachycardia, and send him to see an endocrinologist. Later that week, the endocrinologist sends Bob's blood out for examination; the results suggest that Bob is suffering from acute hyperthyroidism and so the endocrinologist prescribes Bob 30mg each of Methimazole and Propranolol to help combat his symptoms. The first day after taking his new medications while he is turning in for the night, Bob again finds that he is unable to breath, experiences what can only be described as a falling sensation, and decides to call an ambulance. Once in the ambulance, the paramedics begin to take Bob's information. What is Bob's date of birth? Who is Bob's doctor? Is Bob a smoker? What medications is Bob on? Does Bob have any serious medical conditions? Any allergies? Throughout the question and answer period Bob's heart rate rises to 180bpm, close to the suggested maximum rate for his age (40). Once in the hospital (a different hospital from his first visit), the doctors begin to ask Bob the same questions that the paramedics asked and suggest that Bob's Propranolol dosage is much too high. The whole time, Bob is in pain.

Bob now has medical (and financial) records spread throughout both hospitals where he was admitted, the private ambulance company that drove him to the hospital, the commercial lab that processed Bob's blood work, the medical center where Bob saw the endocrinologist, and another medical center where Bob's family doctor, Alice, practices. The next time Bob visits Alice, she asks him to repeat the information that he learned from the other doctors and tests, since the digital files containing specialized information about Bob's condition that did get sent to her office (not all of the files had in fact been sent) were not entirely clear to her (recall that Alice is a general practitioner). Bob may have been spared some time and considerable physical discomfort had the information generated between the various doctors and specialists existed in a single easy-to-use health database rather than existing in data silos each with their own unique terms, systems, and processes. The ambulance drivers had no previous record of Bob's hyperthyroidism or the medication he was on—information that could save lives in more serious cases. Had Bob's data existed (or been aggregated) in a single database, such data would then have to be searched according to a standardized vocabulary. For this to happen, the data (everything from Bob's blood test results to Bob's electrocardiogram) would need to be defined and categorized according to shared meanings and a common taxonomy (a scheme of classification) in a way such that everyone, from the endocrinologist to Alice, would be able to understand how to search for and interpret the data using this new representation.

In the field of computer and information science, a hypothetical taxonomy and its attendant meanings, definitions, and relations is an example of a computational ontology, which is defined by the Italian computer scientist Nicola Guarino and colleagues

(Guarino is one of the earliest practitioners of applied ontology work and a pioneer in the field) as a "means to formally model the structure of a system, i.e., the relevant entities and relations that emerge from its observation, and which are useful to our purposes" (Guarino, Oberle, & Staab, 2009, p. 2). Ontology, then, concerns the shared meanings that are ascribed to particular entities, attributes, and relations.

Ontology-building involves multiple partnerships and is a largely interdisciplinary endeavor (Okada & Smith, 2008). Work in applied ontology-building can involve philosophers who are able to logically define categories and their relationships (Smith, 2003), as well as computer and information scientists who construct technical software for classification systems in fields as diverse as military intelligence (Dragos, 2013), bioinformatics (Bodenreider & Stevens, 2006), and management (Allen & March, 2012). When data are well-structured and annotated correctly, such data can be searched, compared, and interpreted using computational reasoning (Eisinger & Małuszynski, 2005). In bioinformatics, ontologies such as the Gene Ontology are used to process data from experiments to compare multiple items such as diseases and species (Mayor & Robinson, 2014). In business, ontologies can be used to make data available across multiple departments (Michel, 2016). Ontologies include definitions and relationships that are logically formalized using semantics to produce greater data intelligibility (Heiler, 1995). Outside of specific domains, ontologies can enable better understanding among various and diverging research fields by allowing better data communication (Kallinikos, Aaltonen, & Marton, 2013). Ontology construction is now a topic of research in many different areas including bioinformatics (Stevens, 2013), geospatial analysis (Kitchin, 2014), management systems (Orozco, 2012), chemical engineering (Marquardt, Morbach, Wiesner, & Yang, 2010), commerce (Fensel, 1998), judicial knowledge (Casellas, 2011), computer and information science (Poli, Healy, & Kameas, 2010), and even food (Boulos, Yassine, Shirmohammadi, Namahoot, & Brückner, 2015).

Ontologies can reduce bureaucratic bloat (Riaño, 2009) and assist in discovery and innovation using Big Data (Wagner-Pacifici, Mohr, & Breiger, 2015; Kitchin & McArdle, 2016) while enabling better human-computer understanding across complex socio-technological domains (Geels, 2010); they allow for data integration and harmonization so data from divergent domains of application can be synthesized to produce new forms of knowledge (Bodenreider, 2008). Ontologies sit at the intersection of science and technology—it is neither science itself nor technology that explains the relational work of ontologies. By harnessing preexisting scientific data with computational sorting tools, ontologies produce new knowledge that – though not strictly tied to the scientific method – contributes to scientific progress in ways that did not exist previously (Brodaric & Gahegan, 2010). Ontologies fill in the gap that exists between science and technology by using technology to expand preexisting scientific data to produce new scientific discoveries (Bundy, 2008). To use a rather clunky metaphor, ontologies are like a good translator that not only translates languages instantly but also comes up with new compound words, neologisms that are more accurate constructs of words that previously existed in the original language (for example, connecting ‘after’ and ‘noon’ to create ‘afternoon’). Applied ontology uses methods from philosophical ontology to provide comprehensive and accurate representations of data that are used in scientific research (Munn & Smith, 2008).

1.5 A Theory of Informational Ontology and Technologies of Individuation

Currently, there is little literature that considers applied ontologies as communicative technologies or intersecting sites of communicative and philosophical theorizing. I describe computational ontologies as technologies of individuation (Simondon, 2005) that partition reality to translate disparate data and produce informational cohesion and semantic interoperability (Heiler, 1995; Pileggi & Fernandez-Llatas, 2012). In doing so, I argue that computational ontologies re-ontologize (Floridi, 2013) reality, including elements of social reality via social ontology (Ziv & Schmid, 2014; Gallotti & Michael, 2014) and thus that they carry ensuing ethical consequences. Data are metastable entities that can be interpreted, processed, and shaped according to various qualitative and quantitative techniques. By metastable I mean that data are never objective and transparent informational entities but rather are more like diamonds that refract light in multiple ways. Data can be processed differently depending on the techniques and frameworks from which the data are approached. As such, data are never ‘raw’ (Gitelman, 2013) and should not be treated as such in ontology work.

According to the French philosopher Gilbert Simondon (1924-1989), information exists in a state of metastability within a multi-dimensional and preindividual system (data are not completely transparent individual objects) where information can be seen as having the quality of being interoperable and indeterminate, carrying the potential to individuate into a variety of forms (as semantic information, environmental information, biological information, instructional information, and so on). Rather than stop at a definition of information in terms of its probabilistic transmissibility along a channel as given in Claude Shannon’s (1916-2001) mathematical theory of communication

(Simondon heavily engaged the work of early information theory and cybernetics), Simondon offers that information can be thought of in terms of instances where one type of information interacts with another in an event that produces a fundamental change in ontology.

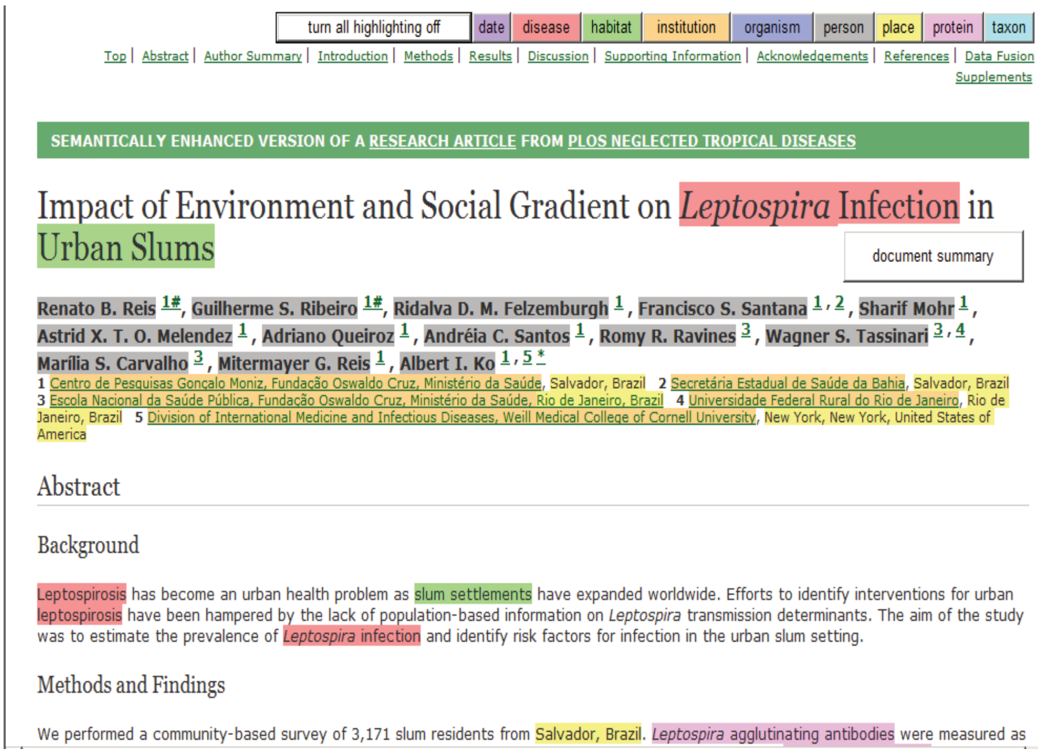
In *Individuation in Light of the Notions of Form and Information* (2005), Simondon wrote that the notion of form must be replaced by that of information and his point was that arguments about dichotomies such as subject/object, realism/antirealism, and abstract/concrete are in some sense resolved once reality is thought of in terms of levels of informational abstraction. The main question that Simondon sought to solve was the following: How is it that entities, attributes, and relations in the world individuate? His answer was that it is in terms of information that agents are able to come to an understanding of how individual entities, attributes, and relations appear in reality. On this theory, doing philosophy in terms of objects and forms is outdated and Simondon saw this right at the beginning of the information revolution.

According to Simondon, information does not exist in a single, homogenous reality—rather, it exists according to ordered levels of abstraction which can disclose new informational levels where entities emerge at greater or lesser degrees of granularity. Information either at the unit or transindividual level (transindividual in the sense of information's capacity to represent multiple things at once) is understood here as not only deposited in a form that is given (specific, well-formed data) but instead is the potential communication between disparate realities that exist but which become actualized through information. Data are not only about entities but also represent potential meanings that can arise when a discovery of individuation activates communication between virtually disparate dimensions that form a system of information.

As technologies of individuation, computational ontologies translate information from a state of metastability to stability—by partitioning reality according to levels of informational abstraction, computational ontologies can constrain wildly divergent data according to logically coherent principles of organization, thus producing semantic interoperability between heterogeneous datasets. In this way, computational ontologies can be contrastively approached as a new form of data processing viewed as the integration of contrasting scientific data into semantically unified wholes. The theory of technologies of individuation acknowledges that even seemingly well-formed data are rarely static entities and that they have no unity or identity because data are never an end in themselves; rather, they require an individuating system. Information as data or message does not provide a complete picture and greater attention from communication researchers should be paid to technologies of individuation like computational ontologies where heterogeneous data are translated and new information is constituted. Simondon viewed information as multimodal and as something that could be exchanged not only between beings who are already individuated but also within systems to come that are productive of new individuations. On this theory, information is internally complex and should not be confused as consisting only of things like media signals.

Informational ontology describes the virtual structures within which technologies of individuation interact to produce something that is ontologically new. Translation indicates the meeting of disparate informational realms and signals the beginning of the process of individuation—it points to the emergence of a new informational structure, one that resolves a disparity between fields that come together to actively produce the potential that lives in matter. Translation signifies domains of potentiality that are represented in the interfacing of information that is inherent to their respective systems, unlocking and reconfiguring one another. For example, apps have altered ontology to produce new cartographic realities, as have global positioning systems and even the postal infrastructure (Bratton, 2016)—such technologies introduce layered informational realities that produce new ontologies.

As technologies of individuation which carry the potential to re-ontologize (Floridi, 2013) reality, computational ontologies treat information not as an ideal or absolute entity but as inhering to materiality in such a way as between parts of a system. Information, rather than acting as bits within a channel, fundamentally alters the system itself, potentially producing new ontological realities by reconfiguring opposing realms in a way that resolves a contradiction. One current example of technologies of individuation can be found in the modern practice of annotating scientific research papers using a standardized ontology—a practical problem. Increasingly, publishers are investigating various methods to enable the tagging of scientific literature in ways designed to make their contents more easily searchable to computers and subsequently open to scientific experimentation. To maximize this potential, a single set of terms is used for tagging the literature in a given domain. The problem to be solved is how to select the set of terms (ontology) for each domain. Once such computational ontologies are introduced, they are the product of not only new ontological technology but also produce ontology and re-ontologize (Floridi, 2013) scientific literature databases to produce new connections and potentialities that did not exist previously. When one scientific article's data can be meaningfully compared to another, such a process stands to alter understanding in the ontology of science.



turn all highlighting off | date | disease | habitat | institution | organism | person | place | protein | taxon

Top | Abstract | Author Summary | Introduction | Methods | Results | Discussion | Supporting Information | Acknowledgements | References | Data Fusion Supplements

SEMANTICALLY ENHANCED VERSION OF A RESEARCH ARTICLE FROM PLOS NEGLECTED TROPICAL DISEASES

Impact of Environment and Social Gradient on *Leptospira* Infection in Urban Slums

document summary

Renato B. Reis [1#], Guilherme S. Ribeiro [1#], Ridalva D. M. Felzemburgh [1], Francisco S. Santana [1,2], Sharif Mohr [1], Astrid X. T. O. Melendez [1], Adriano Queiroz [1], Andréia C. Santos [1], Romy R. Ravines [3], Wagner S. Tassinari [3,4], Marília S. Carvalho [3], Mitermayer G. Reis [1], Albert I. Ko [1,5*]
1 Centro de Pesquisas Gonçalo Moniz, Fundação Oswaldo Cruz, Ministério da Saúde, Salvador, Brazil 2 Secretária Estadual de Saúde da Bahia, Salvador, Brazil 3 Escola Nacional da Saúde Pública, Fundação Oswaldo Cruz, Ministério da Saúde, Rio de Janeiro, Brazil 4 Universidade Federal Rural do Rio de Janeiro, Rio de Janeiro, Brazil 5 Division of International Medicine and Infectious Diseases, Weill Medical College of Cornell University, New York, New York, United States of America

Abstract

Background

Leptospirosis has become an urban health problem as slum settlements have expanded worldwide. Efforts to identify interventions for urban leptospirosis have been hampered by the lack of population-based information on *Leptospira* transmission determinants. The aim of the study was to estimate the prevalence of *Leptospira* infection and identify risk factors for infection in the urban slum setting.

Methods and Findings

We performed a community-based survey of 3,171 slum residents from Salvador, Brazil. *Leptospira* agglutinating antibodies were measured as

Figure 1.1 Tagged Scientific Article (Smith, 2009)

In Figure 1.1, tags include *date*, *disease*, *habitat*, *institution*, *organism*, *person*, *place*, *protein*, and *taxon*. By tagging vast quantities of domain-specific literature so, the ontology can produce new insights by cross-referencing the articles according to these new identifiers. Such ontology engineering facilitates the harmonization of data to produce intelligible content out of data that may not have been previously connected. Computational ontologies are also designed by engineers and computer scientists to aid in knowledge management, data integration, and decision support. For example, an organization may wish to create an ontology of the classes and relations of its employees. *Manager*, *packer*, and *security guard* might represent abstract categories of employee class while *interacts_with*, *cooperates_with*, and *backup_for* might represent categories of relation. The main components of a computational ontology are the entities, attributes, and relations that cohere in a specific domain. Many organizations have data spread over multiple locations and databases that often lack consistency; the data remain unstructured and piled in data silos that are unable to interact to produce a wider scope of the domains in which the data are contained (Roussey, Pinet, Kang, & Corcho, 2011). The main function of an ontology is intelligence growth and the ability to produce a larger picture of organizations (in the broad sense of the word). Ontology user experience is also at the heart of ontology construction in that ontologies must be easy-to-use and friendly for the researcher who is searching through the data (Warren, Mulholland, Collins, & Motta, 2014). Each ontologist must ask: 'How am I to bring disparate data together to create intelligence across organizations for the user?' The integrating and harmonizing of datasets are seen as valuable to science as basic research; combining vast quantities of data across disciplinary domains stands to increase knowledge growth exponentially.

1.6 The Tower of Babel Problem

Like many growing disciplines, ontology research has its own perennial problem. The hard problem of ontology is known as the Tower of Babel problem (Smith, 2004a; Blass, Gurevich, & Hudis, 2007). The Tower of Babel problem (ToB) states that each time a new database is constructed new terms are developed that represent an ever-changing language thus complicating applied ontology-building, the goal of which is to produce ontologies that can last over time. Sometimes a group of databases can be partnered in such a way that they share the same language; however, in the vast majority of cases databases that exist in the same domain remain effectively blocked from each other due to differences of taxonomy—their definitions and relations do not cohere. The ToB problem is what prevents scientists from realizing the full potential of ontologies; each new category and relation from a different domain threatens to undermine an ontology by virtue of the heterogeneity of the data structure. These stem from larger problems concerning distinctions between ontology, epistemology, and terminology. As John Doe (real names of interview subjects withheld to protect anonymity), a leading scientist involved in applied ontology-building explained to me:

> I quite often make the claim that if we want to advance science we should stop talking because I honestly believe that natural language currently did not evolve in a way that allows us to speak about reality in the way that we should speak because of all the discoveries that have been done. And that is not just a matter of adding new terms to the vocabulary. It has mainly to do with the way that when we hear sentences and how they are phrased grammatically that gives us already a bias toward interpretations, things that have clearly been demonstrated in the work by George Lakoff. Sometimes we fall also in that trap; sometimes I read my own work from five years ago and I must say that I am not sure anymore that I actually meant what I said (J. Doe, phone interview, September 9, 2015).

Doe's concern about the nature of language gets to the heart of the ToB problem, however there are also technological and philosophical issues to resolve, including the material constraints for language, such as when species identification changes as scientists learn more about genetic markers. If the goal of ontology-building is to eventually construct a normative hierarchy of entities, attributes, and relations through which human users can process and search large quantities of disparate data, how does the malleability of language and the transient nature of word adoption (particularly in the natural sciences) enable or impede the development of applied ontology? For example, scientists once used the term *luminiferous aether* to refer to the alleged worldly medium that propagated light. Isaac Newton (1642-1726) suggested the existence of such an aether in the third book of his *Opticks* (1718), yet scientists no longer use the term today. This is not only a language problem but also a problem with the materiality that is theorized. Similarly, the *phlogiston theory* once propagated by the German alchemist and physician Johann Joachim Becher (1635-1682) is no longer accepted by science (we now know that what Becher had in mind was the process of oxidation). On an even more basic level, researchers simply tend to use different referencing tags and naming schemes for various kinds of research and data. Given these and other countless examples, how should scientific progress and linguistic impermanence be considered in the practice of applied ontology-building which (requires cohesion), and are there certain approaches to ontology-building that might offer more amenable solutions than others?

One potential answer to ToB has been the linking of datasets (Heath & Bizer, 2011) using hyperlinks; however simple links do not always work and are largely seen as a type of palliative measure. Linked open data is not enough since the data may be incompatible; a link may connect one dataset to another and allow for a connection of datasets but this does not grant the datasets any form of mutual intelligibility or logic (Hepp, 2008). The data must be organized in such a way that definitions and relations cohere among datasets. Links are often notoriously not well-defined and full of redundancies that threaten the potential productivity of data integration. While one database might contain links to another, unless those databases share a rigorously maintained taxonomy of entities and relations such databases will be merely juxtaposed with one another while lacking any concrete opportunity for knowledge synthesis.

Maintaining semantic interoperability (Heiler, 1995; Pileggi & Fernandez-Llatas, 2012) among datasets in the face of data dispersion and the astronomical growth of new data via information and communication technologies (ICTs) has been recognized as a solution to ToB, yet it is incredibly difficult to sustain standards that constrain the definitions and relations of new databases. The outcome is that scientific knowledge growth is siphoned into different domains where data become inaccessible or incomparable (even with the use of a translator). While there has been an increase in scientific knowledge over the last century thanks to ICTs there has been a significant lack of progress in maintaining scientific datasets and data infrastructures in such a way that they remain semantically interoperable with high information quality (Floridi & Illari, 2014). Semantic interoperability is the goal of ontology engineering; rather than allowing science to proliferate into largely diversified repositories of knowledge, ontology engineering seeks to make such repositories open to one another to produce new knowledge.

The ToB problem also has to do with how a single entity, attribute, or relation is to be interpreted when data about said entity, attribute, or relation exists across several sites. Refer back to the case of Bob; information about Bob's hyperthyroidism existed in multiple databases and ontologies. To Alice, Bob's hyperthyroidism manifested as a symptom (his nervousness). To the endocrinologist, Bob's hypothyroidism manifested as the blood test results that indicated incredibly high thyroid hormone levels. The entity in this case is Bob's hyperactive thyroid and information about it has been shared across various databases. Without the information from each database, it is impossible to receive a full picture of Bob’s hyperactive thyroid. Connecting different data points should allow for a fuller picture of Bob’s hyperactive thyroid, but this also says something about the reality of Bob’s hyperactive thyroid. There is an old philosophical joke that might clarify the point. A philosopher puts the following question to her friend: Two humans, a monkey, and a robot are looking at a piece of cheese; what is common to the representational processes in their visual systems? The friend replies: The cheese, of course! Similarly, computational ontologies as technologies of individuation try to recognize that mistakes or contradictions in data categorization are often attributed to a confusion of various levels of perspective and different vantages when attempting to build an ontology around entities, attributes, or relations in reality.

A variety of approaches have attempted to solve ToB and achieve semantic interoperability with virtually all of them sharing in common the distinct feature of adopting methods from philosophy. Ontology has existed as a branch of philosophy for centuries and this has (perhaps not unexpectedly) produced competing views on the subject. If philosophical ontology is abstract, computational ontology is the practical application of specific varieties of abstract philosophical ontology. Much of the literature in computer and information science that deals with ontology engineering has long recognized this debt to philosophy (Zúñiga, 2001), however many do not engage the topic with any philosophical depth (Smith, 2004a). ToB is effectively a philosophical problem as well as a technological and linguistic one that forces ontologists to come up with a way for successfully implementing and practically applying philosophical ontology to computational ontology to make data structures cohere. I asked Dr. Steven

Thomas, one of the most influential ontologists working today (and a classically trained philosopher), if ontology building concerns making different data structures cohere. This was his response:

> I would say that the idea of making data structures cohere can be understood in two ways. One way is that you rebuild the data structures so that they cohere. The problem with that approach is that it's expensive and it creates new errors. Every change can lead to a problem, because somebody makes a mistake or some machine isn't programmed properly. The other way of making data structures cohere is to describe them in a hands-off way, using a common vocabulary.
>
> The common vocabulary raises the following problem, which is that people like to use their own vocabulary. If somebody comes along and they say they know how to build a common vocabulary, which will bring about this trick of making heterogeneous data structures cohere, it's a very difficult one to carry off. You need to know why a given common vocabulary is a better vocabulary than some other proposed common vocabulary. My influence has been in providing principles for building common vocabularies, which are tested in use very thoroughly, over many years now, but also which are well argued for. We have evidence. We have an understanding of why these principles should work. That's one of the reasons why I've been able to have an influence in the way that I have. It's not just that I know that we need a common vocabulary. It's that I also, to some degree anyway, know how to build a common vocabulary in a way that will gain acceptance. That's partly as a result of the philosophical background, but that's not the way to get people to accept something. If you tell them it's philosophy, they will use that as a reason not to accept it (S. Thomas, phone interview, June 11, 2015).

As another well-known scientist put it to me (J. Doe, phone interview, September 9, 2015), we are living in a world where problems of making data compatible and discoverable are becoming ever more urgent. This means that the science that attempts to address issues relating to those problems, which many people in the computer world call ontology, is becoming ever more important. One of the roots of ontology in this new sense is certainly philosophy; the methods of doing ontology work for computer purposes are recognizably philosophical methods. But ontology, as it is now developing, is not a part of science. Rather, just as psychology grew out of philosophy in the nineteenth century as an independent discipline, so too is ontology growing out of philosophy in the twenty-first century, becoming its own independent discipline tied to computer and information science.

1.7 Computational Ontology and Science and Technology Studies

The field of Science and Technology Studies (STS) has produced many volumes on the social, philosophical, and ethical approaches to emerging technologies (specifically ICTs) but has only recently engaged ontology as a theoretical and methodological topic of inquiry (Heur, Leydesdorff, & Wyatt, 2012; Woolgar & Lezaun, 2013). According to Heur, Leydesdorff, and Wyatt (2012), despite this seeming 'ontological turn' in STS there are actually "multiple discussions deploying the language of ontology" (p. 341), including three main themes which include the debate between constructivism and realism, discussion of ontology's instruments and classification, and methodology in the social sciences and the humanities in general. In the debate between social constructivism (Bijker, Hughes, & Pinch, 1987) and technological determinism (Dafoe, 2015), computational ontologies as technologies of individuation fall somewhere in-between.

In *The Cult of Information*, a rarely cited STS polemic from 1986, the historian Theodore Roszak recalls Hans Christian Andersen's children's book *The Emperor's New Clothes*. Early in the text, Roszak offers a response to what he perceives as the growing idolization of information: "Information has taken on the quality of the impalpable, invisible, but plaudit-winning silk from which the emperor's ethereal gown was supposedly spun." The book is an historical overview that analyzes information as a commodity. It argues that information – contrary to the claims of technology enthusiasts at the time – should not be perceived as a causal factor in the shaping of society. Rather, Roszak states that progress is “grounded in the mind's astonishing capacity to create beyond what it intends, beyond what it can foresee.” *The Cult of Information*’s early critique of ICTs and emphasis on their social dimension foreshadowed the trenchant critiques of technology that were to come in the next decade. It also showed, *avant la lettre*, that something called “the social construction of technology” (SCOT) as an academic theory was on the horizon (a Google Ngram search shows use of the term skyrocketing in the late 80s and 90s).

That same year, Langdon Winner published his monumental *The Whale and the Reactor* (1986), which collected many of the celebrated philosopher’s essays on technology and politics. A year later saw the publication of what would become one of the key surveys of this emerging field – *The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology* (1987) – along with a deluge of social constructivist texts on technology in the intervening years. Roszak’s book about information fits squarely into this SCOT tradition. The main underlying assumption of SCOT is that technology is socially embedded and thus that it follows the path of human decisions. A classic example of SCOT can be found in Bijker’s research on the bicycle and the variety of socially related uses it has had throughout history (the penny-farthing was used by young, rich men to display physical prowess and provide a thrill). SCOT, in turn, grew as a reaction to technological determinism (TD), a reductionist view of technology that sees technology as shaping the path of human beliefs and culture. A good example of TD can be found in Lynne White’s famously controversial claim, in *Medieval Technology and Social Change* (1966), that the invention of the stirrup effectively 'caused' feudalism. Similar TD claims can be found in Beniger's (1989) detailed account of how information spurred a new technology revolution in the twentieth century and Friedel's (2007) history of technology as cultural improvement.

Both the SCOT and TD theories are available in strong and weak varieties and both have their critics. Williams and Edge (1996) argue that many researchers "are united by an insistence that the 'black-box' of technology must be opened, to allow the socio-economic patterns embedded in both the content of technologies and the processes of innovation to be exposed and analyzed." Conversely, critics of this tradition have included Winner, who has stated that technologies have their own "politics" and "forms of life." He has also criticized the fact that SCOT offers "no judgment on what it all means, other than to notice that some technological projects succeed and others fail, that new forms of power arise and other forms decline." Books have been published about information in both the TD and SCOT traditions. How computational ontologies as technologies of individuation are to be situated in the debate is not clear, specifically when compounded by ethical questions. Computational ontologies are both socially constructed and depend on previously existing scientific data and technological lifeworlds (Winner, 1986/1988; Ihde, 1990). How, then, do they fit into contemporary TD/SCOT debates?

Whether or not computational ontology suits a TD or SCOT view of the world is likely a futile exercise. Since ICTs and emerging media like computational ontologies are interdisciplinary and social technologies, the TD/SCOT views are perhaps ill-equipped to parse through the changes that such technologies bring. If anything, such technologies are surely not entirely deterministic entities; they require the social to be put to use. On the other hand, such technologies are developing at an exponential rate, and more and more fields today seem to be plugging in to the Ontology of Things. Technologies bend and are shaped by human choices and decisions that are made every day, yet the distinction between what Aristotle called e*pistêmê* (knowledge) and *technê* (craft) might not be as clear as it seems and they are likely two sides of the same coin. Contrary to most historically important STS studies, the important question that should be asked is not a causal one about whether computational ontologies are technologically deterministic or socially shaped but rather a much more pragmatic one about whether or not the changes that are produced by computational ontologies as technologies of individuation are positive or negative ones. The scientific consensus seems to be that computational ontologies are inherently positive technologies that foster scientific collaboration (one ontologist I spoke to said that he could not think of a single ethical problem related to computational ontology), but it is not at all clear that this is the case.

A Google Ngram search for the word "ontology" shows that use of the term increased steadily in the 1950s and skyrocketed after the 1980s (Figure 1.2). "Ontology" was used intermittently in early conversations about AI, information theory, and computer science (McCarthy, 1980) but it was in the early 90s with the publication of a series of papers by Thomas Gruber (1991, 1992, 1993, 1995) that ontology spread as a popular term for achieving semantic interoperability among databases predicated on first-order logic. Previously, terms such as 'knowledge engineering' and 'knowledge representation' were used to describe the work of data integration. In his entry for

"Ontology" in the *Encyclopedia of Database Systems* (2009), Gruber writes that "in practice, the languages of ontologies are closer in expressive power to first-order logic than languages used to model databases" (p. 1963). The languages used to model databases and query them are indeed important to ontology yet what Gruber calls attention to is the purely philosophical first step of ontology construction and how this step must occur before applying the ontology to technological systems, or what amounts to having it manifest in a database.

Figure 1.2 Google Ngram Search Results for "Ontology"

An ontological commitment to some entity, attribute, or relation is typically a first-order logic combined with a specified domain of discourse on top of which quantified variables range (recall Quine's bound variables). Once this has been achieved, *ontology representation languages* (ORL) are constructed, which can be seen as the practical application of first-order logics using technological vocabularies; these include description logics (knowledge representation languages) such as F-logic, "a formalism that integrates logic with object-oriented programming in a clean and declarative fashion" (Kifer, 2005, p. 22), the Resource Description Framework (RDF), a representation system for sharing knowledge and data on the web, and more expressive ORLs such as the widely-used Web Ontology Language (OWL) and Dublin Core.

Ontology engineering also requires the use of specific *methodologies* that are relatable to software engineering given that both involve the use of diverse data in distributed settings; one such methodology is Distributed Engineering of Ontologies (DILIGENT) which seeks to avoid the problem of peer-to-peer systems being inadequately supported by a centralized ontology by supporting "domain experts in a distributed setting to engineer and evolve ontologies" (Pinto, Staab, Tempich, & Sure, 2006). There are tools such as Formal Concept Analysis and OntoClean which help ontologists grandfather old ontologies into new ones, as well as methods for designing, learning, and evaluating ontologies. There are the *computational ontologies* themselves, entities such as the popular Descriptive Ontology for Linguistic and Cognitive Engineering (DOLCE), Core Ontology for Multimedia Annotation (COMM), and Process Specification Language (PSL). Lastly, there are *upper level ontologies* which can be considered ontologies of ontologies; the most widely used upper level ontology is the BFO. BFO is one of the most successful ontologies in the world and its creator, Smith, is a professionally trained philosopher.

Ontologies need infrastructures to sustain themselves and the scalability of ontology engineering is important for international ontology growth in multiple domains. Managing multiple ontologies in indexes has become a key area of research to enable the easy searching and reuse of ontologies. Ontology repositories such as the Open Biological and Biomedical Ontologies (OBO) Foundry provide tools for supporting such searching and reuse. The OBO Foundry is the world's largest cohesive ontology repository and is organized using BFO principles. Ontology infrastructures like the BFO are rarely studied, yet as data infrastructures that enable scientific research they are central to the communication of some of the world's most advanced scientific data. Ontology infrastructures are data assemblages that can be analyzed at various levels of abstraction (Floridi, 2011) and thus necessitate a mixed methods approach to capture the multiple perspectives from which scientific data infrastructures can be viewed, including technical, personal, and ethical dimensions.

Technologies of individuation have the ability to re-ontologize the world. While not entirely in line with the SCOT approach to technological development, technologies of individuation are not fully deterministic and are often used according to the will of powerful institutions that control them (increasingly, intelligence agencies). On the other hand, as computational ontology research progresses, it will be important to pay attention to how computational ontologies engender more or less political attitudes among the groups who use them—and what the presence and absence of such articulated politics means. Data that are manipulated by technologies of individuation are not only unique theoretical objects that can furnish new scientific understanding; they can open profound philosophical and ethical questions that deserve to be asked. Computational ontologies change ontological understanding of the world by defining what that world is made up of. In doing so, scientific discovery should not trump political justice and data science should not put progress over people. What, for example, to make of the Military Ontology and similar intelligence ontology projects that are being carried out to individuate the reality of war (and the individuals involved) into computationally processed entities? I reserve such ethical questions for the final section of this dissertation. Sidestepping decades-old debates between constructivism/realism and ontological methodologies in the humanities, my intention is to focus on the logic, instruments, groups, institutions, science, and classification of computational ontology. In doing so, I foreground the philosophical, historical, social, instrumental, and material dimensions involved in the making of computational ontologies as technologies of individuation to reach some of their more ethical consequences (Hacking, 1983; Haraway, 1991; Latour, 1988).

1.8 Informational Ontology as a Realism of Relations

Gruber (1995) is credited with defining an ontology as an "explicit specification of a conceptualization" (p. 908) and today this remains the most popular definition for what a computational ontology is. The first step towards achieving semantic interoperability between datasets is achieved through ontology builders coming to agreement on shared conceptualizations for entities, attributes, or relations and the second step involves specifying the data according to these shared conceptualizations (Gruber's simple and widely-used definition gets at the heart of this process). Concepts have by and large been the preferred way to do ontology work, especially among computer and information scientists.

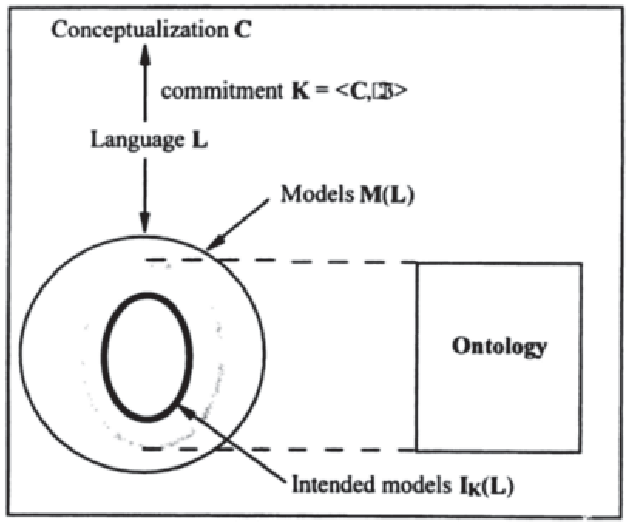


Figure 1.3 Ontology as a Specification of a Conceptualization (Guarino, 1998a)

In his introduction to the proceedings of the first international conference on Formal Ontology in Information Systems (FOIS), the Italian engineer and ontology pioneer Nicola Guarino (1998a, 1998b) indicates that the word *conceptualization* should be borrowed from AI and used to indicate the philosophical notion of *meaning* so that two ontologies can be different in vocabulary while maintaining a commitment to the same entity (Figure 1.3). On Guarino's interpretation, that conceptualization is meant to be language-independent is trivial (English and French words can share the same referent), yet the term raises problems with respect to the arbitrary nature of its claim on the reality or non-reality of entities. A shared conceptualization in the applied ontology sense implies something that originates by method of an arbitrary *choice* independent of reality, but such criteria would not be strong enough when attempting to build ontologies for use in the natural sciences which demand that ontologies represent scientific laws and structures accurately as they exist in reality and pointing to specific referents. A commitment to concepts in ontology-building can lead to a number of infelicities arising from a lack of commitment to real entities. Concepts can change at any time given group consensus and suggest a shared arbitrary choice while a realism-based approach leans on the permanence and reality of scientific theories; realism-based approaches attempt to accurately describe entities to establish regulatory permanence and longevity across diverse datasets—such is the value of an applied ontology.

Building on the work of ontological realists involved in the practice of applied ontology-making, I argue that realism-based ontologies offer a better solution to ToB in place of concept-based ontologies while emphasizing the important role of *information* in such realism-based approaches. For example, information can exist *as* reality, *for* reality, and *about* reality (Floridi, 2011)—distinctions that matter when attempting to label artifacts represented by data. A crude example might help to elucidate this point. Blog posts are often indexed via the use of tagging mechanisms to create a folksonomy. A blog may host many posts that are organized according to conceptual tags; a post containing a picture of a dog jumping into a lake can be tagged with such common nouns as *picture*, *dog*, *lake*, etc. But the picture can also be tagged as *45332249*, *Rex jumping*, *excited by a bee sting*, or *catching fish*. Posts are given names and categorized according to the individual users who upload them (though other users may upload photographs using preexisting tags if they wish). Of course, there are many photos that align nicely into categories; millions of posts use the label *cat* or *Thanksgiving* to label their photographs. However, there are also countless posts and photos that are labeled conceptually and in many cases these make no sense to other users. One user may upload a photograph of their high school and label it *nightmare*. Now, that poor individual's high school may well have been a nightmare for them, however for the purposes of rigorous scientific categorization such a concept-based taxonomy does not suffice.

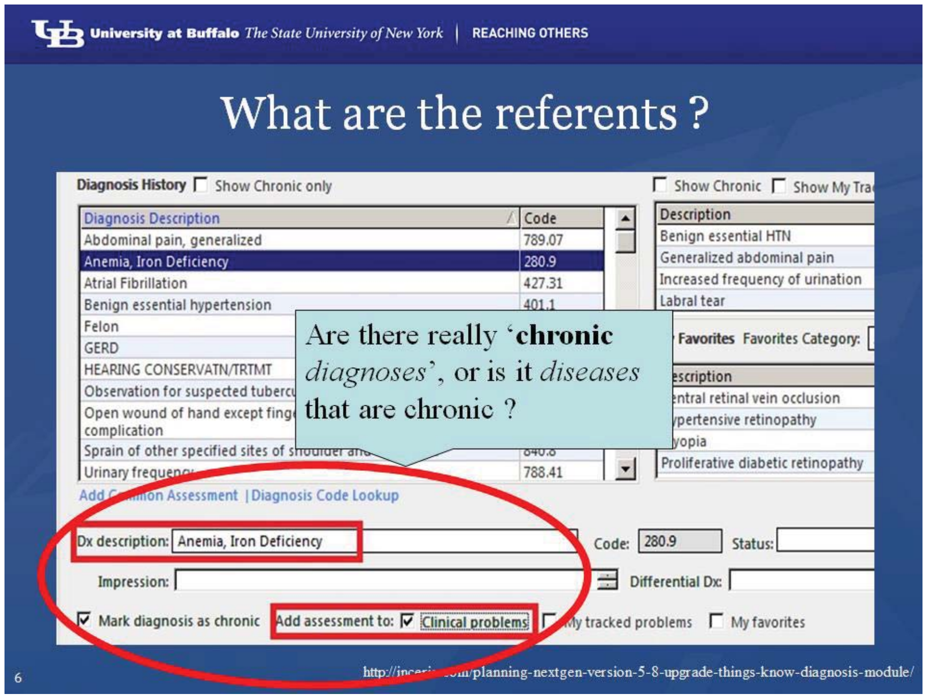


Figure 1.4 Example of Referent Tracking Confusion (Ceusters, 2015)

Tagging photographs online may seem trivial when making a case for a realism-based approach to applied ontology-building, yet the importance of correctly identifying features of reality is extremely important in the natural sciences. The practice of *referent tracking* – the goal of which is to “create an ever-growing pool of data relating to the entities existing in concrete spatiotemporal reality” (Ceusters & Smith, 2007) – in the health sciences works the same way as standard tags that are used online; scientists must come up with ways to track certain entities over time and space with the eventual aim of including these entities in a shared ontology. Representing reality is not easy, as noted with genetic markers as the ontological space for species identification versus morphological features that were once used—both are material. Ceusters (2015) shows a standard form used by doctors in medical settings with a check box that has the words "Mark diagnosis as chronic" next to it; he then asks if there are *chronic diagnoses* or if it is *diseases* that are chronic (Figure 1.4). Further complicating the matter are electronic records such as x-ray and scan images, receipts, prescriptions, patient information forms, etc. Such seemingly trivial errors in specification in these areas can lead to huge problems when it comes to sharing data between databases and among various users.

Computational ontologists generally hold that concepts make up ontologies for knowledge representation, yet the term 'concept' is not often well-defined by ontology engineers in computer and information science. Though many agree that concepts are the product of human cognition, there is some disagreement about how this applies to the realm of ontology building. Following Smith (2004b), I defend the notion that ontologies developed for scientific research should be "understood as having as their subject matter, not concepts, but rather the universals and particulars which exist in reality and are captured in scientific laws" (p. 73). Simply put, such scientific laws represent reality as captured in universals. Such a realism-based approach to ontology construction yields rigorous formal axioms to describe the relations and entities that exist in many influential ontologies. Against such a view, Smith suggests that proponents of various types of idealism might argue that (a) there is ultimately no objective reality to which concepts might correspond, (b) we can never know objective reality so there is no point in trying to establish correspondence, and (c) reality is nothing more than concepts.

In much of the scientific literature, the conceptual/idealist position prevails. However, it can easily be shown that the reason for this is the fact that such endorsements do not take into consideration levels of abstraction, the distinctions that are available after considering time/space, and the differences between entities and relations in reality or names, concepts, or descriptions in terms of language. Take the following example from Smith, the staunchest defender of ontological realism. In linguistics, the assertion *is_a* is an assertion about meaning. A sentence such as *lytic vacuole is_a vacuole* (vacuoles are bubbles in cells) is not an assertion about lytic vacuoles; rather it is an assertion about *language use*. It tells us that the meaning associated with the name *lytic vacuole* is narrower or more specific than the meaning associated with the name *vacuole* by this or that group of subjects (2004b).

Smith claims that there is a certain way in which linguistic and syntactic mechanisms at work in ontologies lead ontology engineers into thinking that they are talking about separate concepts when in reality they are talking about specific linguistic levels or types/sets of syntactic organization. The linguistic or syntactic element however does not constitute a separate concept tied to a real entity (this might explain the confusion that all entities must be referred to as concepts since there are different ways of referring to entities). This does not mean that entities do not exist. The fallacy lies in suggesting that there is no entity or only concepts given that, on the conceptualist's theory, everything can be reduced to concepts; however most concepts can be attributed to linguistic, syntactic, informational errors of leveling. In engineering there is a similar confusion; when engineers speak of conceptual modeling they often think that they are offering only data or information about an entity and that is all. But upon closer inspection the actual practice of modeling involves building models of reality once a specific ontological commitment has been made and a certain level of informational abstraction has been adopted. Once this happens, modeling does not *only* represent a concept or mere data or information—the model makes a claim on a part of reality and thus changes it (think of DNA sequencing).

Ontologies as representative models of reality do work in the world. Yet many practicing ontologists still believe that concepts provide the content of ontologies. Ontologies are generative artifacts that facilitate knowledge synthesis via the practice of semantic interoperability but require real referents. One ontologist explained this to me using the following example. In linguistics, theories and definitions have been developed that treat the basis of terminology as consisting of concepts. On these terms, most of the things we talk about are concepts; concepts are everything and anything that is conceivable. Thinking of ontology in the same terms, he says, is a mistake:

> Ontologies should not contain what is conceivable. They should contain what existed or has not existed. Now, the fact that you can conceive of something or that people can conceive of something; that should go in an ontology. So ‘people who believe in the devil,’ you can say in one way or another, that’s an acceptable defined class in a realism-based ontology, but ‘devil’ itself should not be there (J. Doe, phone interview, September 9, 2015).

This distinction regarding treating concepts as the either the *content* of ontologies or as the *ideas* that individuals have (which can be realistically represented in an ontology) is important in that it helps to identify mistakes in ontology categorization. It also points to the importance of stressing uniform realism in ontologies that must integrate divergent datasets. Popular description logics that have been used in ontology engineering since the 1990s (a family of formal knowledge representation languages) and similar tools are good in detecting logical inconsistencies in representations. They are able to do classifications and then show when the system comes up with a certain classification and whether it is an intended model or not. Yet, such logics still allow assertions about generic reality that are actually not true. There is no description logic that is able to tell that there are no unicorns or that unicorns are not horses with one horn on their foreheads. Such terms can be included in description logic, which has always been understood as being concept-based. The advantage of *ontological realism*, according to an ontologist I spoke to, “is that you do not look through the glasses of language or the glasses of concepts to see what is there. But you try to purely reason in first-order reality itself, or what should be there if such and such is the case.” They continued:

> Of course you can't do that without observation and without interpretation. For that reason it’s extremely important that you keep a distinction between what information and assertions are and how they are formed. How do you relate them to what they are information about? That has been the constant search that I have been doing (J. Doe, phone interview, September 9, 2015).

Concepts, the ontologist held, are not accurate reflections of the information that data are supposed to be about. To come to a more accurate representation of entities in reality, ontologies require informational accuracy about real entities. Another ontologist echoes a similar sentiment:

> There are two schools of thought in ontology, simplifying a lot. One school says we can build ontologies using pattern recognition applied to free text or applied to word lists. The other school says that we have to build ontologies by using principles and by manual effort, largely. I belong very firmly to the second school of thought, and if I'm challenged I can demonstrate that the manually and theoretically, soundly built ontologies are still way better than those ontologies which are created by machine extraction (S. Thomas, phone interview, June 11, 2015).

The traditional way of doing ontology work using description logics and machine learning is largely the product of the conceptualist way of thinking about language and ontology and has prevailed since the 1990s. The realism-based efforts of ontologists such as Ceusters and Smith revolutionizes the field of applied ontology by treating computational ontologies as technologies of individuation that restrain what can be referred to in an ontology to manually construct ontologies that are logically sound and based on universal structures.

The practice of applied ontology-building necessitates the careful consideration of levels of informational abstraction. As a standard philosophical principle, levelism has enjoyed a long and fruitful career (Poli, 2001); the notion of thinking in terms of different levels has existed since antiquity, for example in Pyrrho and the beginning of skepticism or Plotinus and the Neoplatonic synthesis. Today, types of levelism are used in the study of AI and computer science, often with special attention paid to the modeling of these structures from a variety of perspectives. Floridi is one thinker in a small group that is leading this effort with his own unique brand of *epistemological* levelism. Levelism has become a useful conceptual tool to help clear up thinking about complex technological problems like ontology-building. Understood as a methodology, Floridi provides a comprehensive account of what he calls *the method of levels of abstraction* (LoA) that can be used in the analysis and design of ontologies. Distinguishing three general types of data, including information *as* reality, information *for* reality, and information *about* reality, Floridi offers that the method of LoA can clarify some of the inconsistencies that arise in ontology work.

Informational ontology is a *realtional* ontology that does not suffer from abstract conceptualizations and is predicated instead on the relations that obtain between real objects as they are informationally represented according to scientific laws. It follows the notion that ontologies are not merely concepts of entities, attributes, and relations but that an ontology is a representation of real objects or relations according to accurate knowledge produced by scientific laws and the scientific method. The worry in ontology-building should be that information about entities, attributes, and relations be treated accurately as really existing relational entities themselves (what Floridi has called informational objects) and that these informational entities can be acted upon to produce more accurate applied ontologies. A version of this philosophical thesis has been called informational structural realism (Floridi, 2008a) and it is the philosophy that informs the realism-based ontology in this dissertation. By endorsing informational structural realism (ISR), ontology can bypass philosophical dead ends about realism or antirealism of entities, attributes, and relations to focus more closely on the differences that occur when adopting a specific level of informational abstraction to model reality. In this way, ontology engineering can be refined according to strict informational criteria concerning what might be admissible in an ontology. In the same way that better semantic constraints on an ontology make for greater semantic interoperability across ontologies, understanding ontologies as constrained levels of abstraction (LoA) allows for them to be seen as representing *multiple levels of reality* that do not exclude one another. These constraints actually act as *constraining affordances*; by semantically constraining ontologies at different LoAs, researchers can begin the task of solving the ToB problem by integrating cohesive datasets.

ISR offers a modern solution to longstanding realism versus antirealism debates by arguing in favor of a version of *structural realism* (Chakravartty, 2007) where structures are understood as informational since it is by adopting a level of informational abstraction that we interact with the world. At bottom, ISR states that a theory commits itself ontologically by adopting a specific LoA. Applying a LoA commits a theory to a particular model of a system; by adopting a LoA, the theory decides (much in the same way as Quine's ontological commitment) what kinds of observables should be included in the model. Here, one can understand an applied ontology as a type of model that is generated by a LoA for the purposes of identifying a structure that is attributed to a system in reality. Floridi (2011) has offered the best formulation of this in his SLMS scheme (Figure 1.5). The LoA method informs realism-based computational ontology.

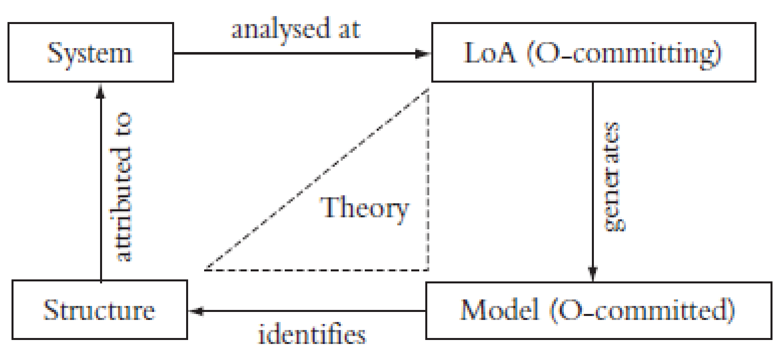


Figure 1.5 Floridi's SLMS Scheme with Ontological Commitment (2008b, 2011)

LoAs suggest a realism of relations (Santos, 2015) in that they treat a minimal commitment to a real entity and the ensuing theory about that entity as a real relational structure that can then be added to using multiple additional levels of abstraction. To use another crude metaphor, one can think of LoAs in terms of a needle that weaves a thread through a piece of fabric. If the fabric represents reality, the first time the needle pierces the fabric and is pulled through until the thread's knot catches can be seen as the first level of minimal ontological commitment that "grounds" the thread, with each subsequent piercing or weave representing another level of abstraction. Cut the section near the initial thread from its knot and the structure is destroyed (think of when you find a loose thread in a piece of clothing and pull on it, thus ruining the garment). Such an understanding of ontological commitment requires viewing the ontology as consisting of *real* and *relational* entities. The language of *concepts* does not fit the types of ontological commitment that can be described using a LoA approach since LoAs are based on some minimally restraining entity that is committed to that *must actually exist*. From there, the ontology can be built up with greater levels of second and third order ontological commitments (and can indeed then include things such as concepts as when we say that a person has "an idea" of something) at more abstract levels of granularity, but these are always based on some commitment to a real entity that grounds the multiple layers of informational abstractions. To emphasize this point, Floridi provides a diagram (Figure 1.6) of ordered ontological commitment (2011).

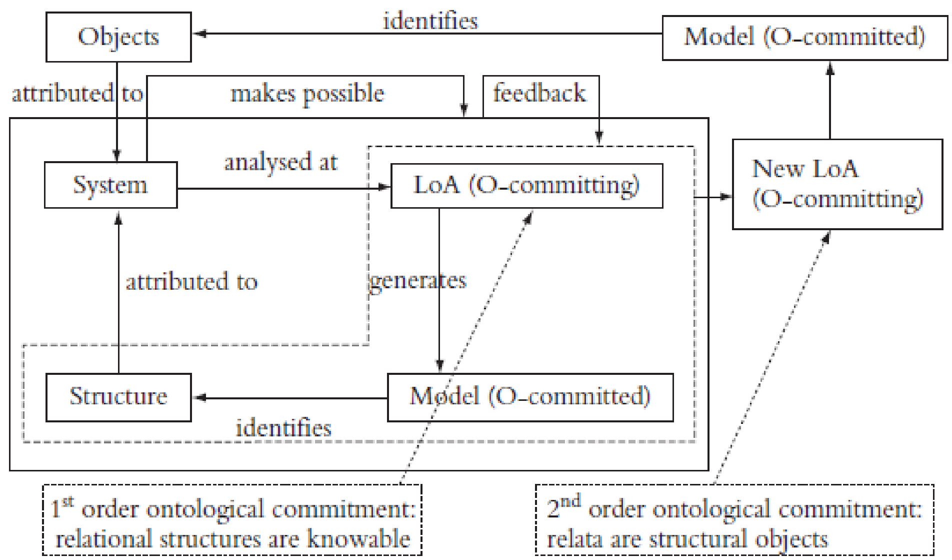


Figure 1.6 Floridi's SLMS Scheme with Ordered Ontological Commitment (2011)

The benefit of this approach shows that there must be some commitment to a real entity before abstracting to greater LoAs that provide views of greater of lesser levels of granularity while also showing that entities do not have to exist at the exclusion of other entities. A LoA acts as a transparent partition of a single reality (Figure 1.7). For example, a folk biology (remembering that *folk* merely implies some acknowledgement of regular things that actually exist) might dictate that there are such things as animals which include such things as fish or that there are animals which include such things as birds which include such things as canaries. There is also such a thing as DNA. Philosophers often enjoy arguing which one is more "real" than the other (the canary or the DNA) however if we view such categories as transparent partitions that exist in reality (LoAs) then we can observe, in a grid-like fashion, that by zooming in or out we do not need to eliminate any of these entities in favor of another, as long as they are grounded in reality and adhere to universal laws and definitions of science.

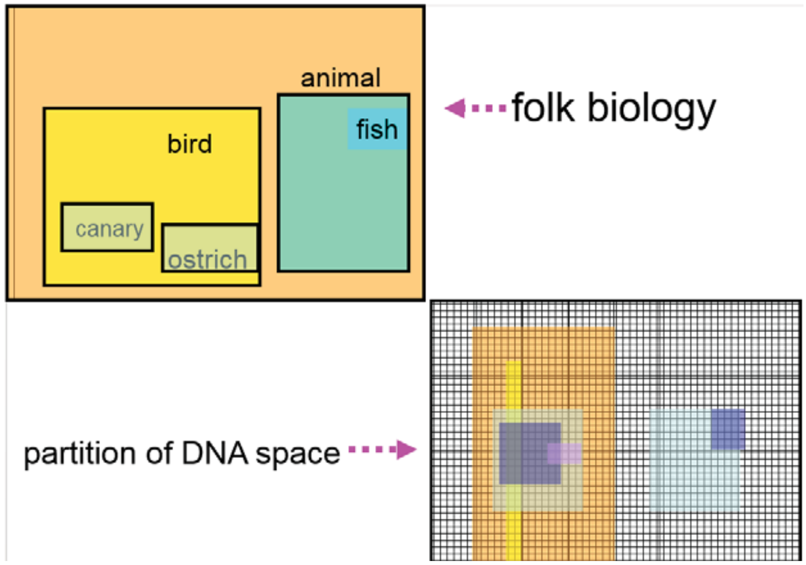


Figure 1.7 Transparent Partitions of One and the Same Reality (Smith, 2009)

Ontologies are rarely prefaced by an adjective unless that adjective describes what is contained in the ontology (gene ontology, blood ontology, etc.). By referring to informational ontology I mean to signify the particular way of doing ontology work that follows a LoA approach, one that is based on ISR. Such realism-based ontology has a long history (dating back to Aristotle) and versions of ontological realism have been refined by various philosophers before being instantiated by scientists and engineers in applied computational ontology. Some of the most successful ontology projects in the world adhere to a strict realism program (Smith et al., 2007).

The revolution currently underway in applied ontology work is associated with ontological realism and the work of ontologists such as Smith and Ceusters and the advancement of this position is important for expanding clarity in the field of applied computational ontology work, including how it is taken up by qualitative researchers, however it also necessitates an analysis of the ethical consequences of such a realism-based method. Any new technology that claims to be based on realism with the power to alter what counts as real needs to be ethically analyzed, especially if those technologies interact with data concerning humans and their environments. Further, as technologies of individuation, computational ontologies are the product of diverse teams and researchers who build them and thus exert forms of influence, power, and control over other social groups. As such, a critical analysis of computational ontology should include analysis of the histories of the groups and individuals involved in building the ontology. Doing this entails interviewing the individuals responsible for ontological realism and its method while also laying bare its institutions, groups, and organizations that have historically contributed to the development of the realism-based method.

1.9 Method and Framework

Ontological realism is important to work in the advancement of interdisciplinary scientific research. As such, applied computational ontologies that use the method of ontological realism should be open to qualitative, critical, and ethical, studies on the impact of ontology on people and their social lives. Following approaches taken in science studies, including studies in Science, Technology, and Society, Science Communication, Rhetoric of Science, Critical Data Studies, and allied fields, this dissertation includes elements of philosophical theorizing, communication studies, and ethnographic fieldwork. Six ontologists involved in ontology research at the National Center for Ontological Research were interviewed over an internet telephone for one hour. The interviews were conducted using Skype and recorded using recorded using CallGraph Skype Recorder for Windows. The interviews were then anonymized and transcribed using CastingWords. Names of interview subjects have been changed to protect anonymity. The series of twenty-one interview questions is located in Appendix A. Each interviewee was asked whether or not they viewed ontology as a communication problem—virtually all replied in the affirmative. Some of the questions asked during these unstructured long form interviews were: How is applied ontology practiced? How do philosophers contribute to building applied ontology frameworks? Why are some computational ontologies preferred over others? Each of the interviews was conducted with the aim of understanding more about how philosophical ontology informs scientific ontology; they also provided a firsthand account of the history of scientific ontology engineering, including the names, places, and technologies that have contributed to these fields. As such, the dissertation follows a mixed methods approach (Creswell, 2013) utilizing a primary mixture of ethnography (Fetterman, 2009) and digital methods (Rogers, 2013). I conduct an ethnography of the BFO as a relational and ecological infrastructure (Star, 1999) via thick description (Geertz, 1973) and long form interviews (Weiss, 1995) with scientists and philosophers involved in the ontology-building process and utilize archival research methods (Ramsey, Sharer, L'Eplattenier, Mastrangelo, 2009; Carmichael, 2012) for probing the large online repository of resources made available by the BFO.

CHAPTER 2. PHILOSOPHICAL ONTOLOGY

2.1 Chapter Summary

This chapter presents a variety of historical philosophies of ontology that feature significantly in modern applied ontology research, specifically in the BFO. It offers some background on early scientific researchers who had interests in philosophy, explains how Aristotle's categories, Husserl's formal ontology, and Ingarden's ontology of time are relevant to some of the applied ontologists I interviewed, and shows why applied ontology should embrace the realism-based method. The main purpose is to show how ontological realism in applied computational ontology grew out of a specific branch of philosophy to help reframe the way that applies ontology work was being conducted in computer and information science. The philosophical theories presented here set the stage for the analysis of the BFO that is presented in the fourth chapter.

2.2 Philosophy's Influence on Applied Ontology

Today there are numerous calls for more participation by philosophers in ontology engineering; Merrill (2011) encourages philosophers to work with scientists to combine the practices of philosophical and scientific ontology. He proposes a strong participation model where applied ontology is not passively appropriated from classic philosophical texts but rather actively discussed between philosophers and scientists so that contemporary philosophers can contribute to modern scientific methods while also progressing philosophy by adapting philosophical methods according to new scientific discoveries. Merrill envisions this interaction between philosophers and scientists as a mutually beneficial relation where trained philosophical ontologists ensure that scientific ontology engineering is both adequate and correct by checking it against formal philosophical logic while learning about advances in science. There are a number of practicing philosophers who might advocate the notion that modern philosophy should concern itself with applied ontology, including philosophers such as Smith, Luciano Floridi, Brian Cantwell Smith, Robert Arp, and Andrew Spear, some of whom are currently engaged in the practice of assisting with ontological engineering (Smith, 1996; Floridi, 2011; Arp, Smith, & Spear, 2015).

After asking an ontologist about the distinction of ontology in philosophy and computing, I followed up by seeing if they could provide me with the names of other philosophers who he could think of as having been similarly involved in philosophical and computational ontology. They explained to me that ontology-building is an amazingly promising opportunity for philosophers to do new and interesting work (which was also well paid work) with a lot of social relevance. However, one of their biggest disappointments has been the degree to which few contemporary philosophers have taken a practical and real interest in applied ontology. There are philosophers who are disposed in a friendly manner towards applied ontology work but there are few who actually do it. Probably the most famous applied ontologist who has ties to philosophy is Patrick Hayes. Hayes, who has a PhD in philosophy, worked as a philosopher for a brief period but then became a computer scientist specializing in AI. He is now one of the leaders in ontology-based logic work in computer science after spending a long career considering philosophical problems from the standpoint of AI (Hayes & McCarthy, 1969). The ontologist went on to say that many of their students have positions now in industry or government and that they are still doing work which is related to the work they did when they were studying philosophy, but that they do not think of themselves as philosophers. I asked an ontologist about what makes for a good ontologist. This was their answer:

> I wish I knew the recipe for creating a good ontology creator. I tried various strategies. One strategy is to train philosophers, and I've been successful in that, but not in every case. The scenario which often happens is I train a philosopher to become an ontologist, but the philosopher still wants to be a philosopher, and they don't want to do this applied work, because it doesn't have the cache of doing “real” philosophy. Still, I have had success in training philosophers to become ontologists. The majority of the people who are doing good ontology work, or what I see as good ontology work within my orbit, are people who are at the fringes of science and computing. They are scientists who use computers a lot in their scientific work, and who see the need for ontology, and who have... They're able to use their computational skills, and their scientific knowledge, in order to build good ontologies. They are the two primary groups. There are some other kinds of people who are very successful at building ontologies, but they are the two primary groups (S. Thomas, phone interview, June 11, 2015).

Philosophers and scientists who understand the importance of computation for developing new scientific techniques tend to be the best ontologists. Such philosophers must have a thorough grounding in philosophical ontology and its relation to natural science.

The ontologies of philosophers such as Aristotle, Edmund Husserl (1859-1938), and Roman Ingarden (1893-1970) are in some sense comparable to the scientific taxonomies proposed by figures such as Carl Linnaeus (1707-1778) in biology and John Dalton (1766-1844) in chemistry, though they are much more general and abstract in nature. Some philosophers are realists about universals, transcending labels such as substantialist (the notion that there is some sort of substantial material reality underlying all things) and fluxist (the notion that reality is in essence comprised of flux) in favor of the realism of categories of such things like objects, events, and processes. The biologist Joseph Henry Woodger (1894-1981) contributed to the beginnings of a realism-based approach to applied ontology work in biology. Woodger was a biologist but he also had interests in philosophy; he translated the work of Alfred Tarski (1901-1983), specifically Logic, Semantics, Metamathematics (1956), a collection of Tarski's work from the years 1923 to 1938. It is through this translation work that Woodger is primarily known to philosophers but he actually wrote books on the ontology of medicine. Woodger's key contribution to founding the field of biological ontology came in the form of his 1937 book The Axiomatic Method in Biology; in it he applied logical axioms from Whitehead's and Bertrand Russell's (1872-1970) Principia Mathematica to medicine and biology. One ontologist told me that Woodger was already in the 1930s doing what they were doing now; using philosophical logic to represent reality in ways which were supposed to be practically useful (Woodger was not necessarily an influence on all ontology work as some ontologists described to me that they had discovered him well after they had already started down the path of applied ontology).

Through the discipline of history of philosophy it has become increasingly apparent that good work in philosophical ontology has been conducted in Austria, Hungary, and Poland. Philosophers in these countries have contributed to philosophical ontology yet some of their works in that tradition were not taken up by philosophers at large due to a general preferred focus in the twentieth century on interests related to either the phenomenological or the existentialist tradition, or by a rather narrow language-focused analytic philosophy. One ontologist told me that the work being done in those regions was a way of doing philosophy of real problems, a real-world focused philosophy (as much continental philosophy is), but one that still had an analytic rigor. It was a combination of the two. For example, the inventor of speech act theory was a student of Husserl by the name of Adolf Reinach (1883-1917). Smith spent time working on Reinach and eventually became interested in Ludwig Wittgenstein's (1889-1951) Tractatus Logico-Philosophicus (1921) because of the connections between Wittgenstein and Reinach's theory of states of affairs presented in Die apriorischen Grundlagen des bürgerlichen Rechtes (1913). In the period towards the end of the twentieth century, Smith saw that some of his work on ontology was being used by computer scientists, particularly in Italy, which was and still is a center on applied ontological research. Yet, computer and information scientists largely ignore the history of philosophical ontology, often preferring instead to include a brief reference concerning their indebtedness to Aristotle. The scientists I spoke to, some of them securing millions upon millions of dollars in research grants and funding, explained to me that the level of indebtedness to philosophical ontology goes much further, through Aristotle and up to figures of twentieth century philosophy—particularly Husserl, and Ingarden. Concepts such as universals and particulars, formal ontology, ontological commitment, and ontological realism – terms used in much of the applied ontology work that is being carried out today – each owe their existence to work that has been done in philosophy. What are some of the important principles and axioms that have to be accepted to commit to the existence of certain entities?—philosophers are the specialists when it comes to such questions.

2.3 Realism, Conceptualism, Nominalism

When I asked one scientist about what they would like to see in a dissertation about ontology their recommendation was to include a comment on an important philosophical difference—in considering universals and particulars, there are philosophical differences between realism, conceptualism, and nominalism. Briefly, realism (sometimes realism is divided between exaggerated realism and moderate realism) holds that universals are real and that they are reflected in the scientific laws of nature. For example, realism would hold that a genus is as real as the species that fall under its domain. Conceptualism holds that there might be universals that individuals can conceive but that those individuals can have no confirmation of whether their universal concepts have any foundation outside of their cognitive processes. Nominalism denies the existence of universals and holds that cognitive processes cannot realize them. Now, some philosophers interested in applied ontology, such as Cocchiarella (2007) and Merrill (2010), hold that it is possible to maintain a blend of realism and conceptualism in something they refer to as conceptual realism. The argument goes something like this. Universals can be predicated of things. In nominalism, there is no such thing as a universal properly speaking; the predication in nominalism is merely a linguistic relation between subject and predicate. For example, in the sentence “the kids may have started the game,” the predicate in the verb phrase “may have started the game” is related to the subject in the noun phrase “the kids” only in terms of the phrase structure of the declarative sentence. There is no universal that is predicated of “the kids” or “may have started the game.” As such, nominalism is a poor choice for ontology-building, which requires some admission of universals that are predicated of things. The conceptual realist might hold that there is an argument to be made in favor of viewing concepts (in place of universals) as the connective tissue that allows for sentences to be meaningful, rather than a mere string of words, as in the case of nominalism. In this vein, Cocchiarella (2007) argues that the "objectivity of referential and predicable concepts consists in their being intersubjectively realizable cognitive capacities that enable us to think and communicate with one another" and that "it is the complementarity between predicable and referential concepts that underlies the mental chemistry of language and thought" (p. 143). Real concepts, then, are found in intersubjective cognitive capacities and complementarity between predicates and subjects. Along these philosophical lines, Merrill (2010) holds that a conceptualist approach to applied ontology-building can achieve as high a standard as realism-based approaches, and to support this he references the work concept-oriented applied ontologists such as Guha and Lenat (1990) and Lenat and Guha (1990). The trouble arises when one recognizes that such conceptualist approaches to ontology-building are rather outdated and used less frequently today due to technical errors. While Merrill would like to critique the realism-based approach in favor of a concept-based approach, the references to applied ontology work that he offers (the work of Guha and Lenat) refer to old engineering literature and ontologies that are no longer in use (for example, the Cyc ontology project from the 90s) and as such he fails to provide a thorough outline of what a new conceptualist approach would look like. The Cyc project in particular – a knowledge representation system design for AI – generated much controversy (Bertino, Piero & Zarri, 2001, p. 275), including problems related to the complexity of the system, unsatisfactory accounts of the concept of substance and of intrinsic versus extrinsic properties, and lacunas in the ontology of ordinary objects.

Building on previous work in the realism-based approach to ontology-building (Smith & Ceusters, 2010; Floridi, 2011), I offer that concepts are primarily either subjective, used to describe language use, or used to refer to formal entities such as numbers. They are assertions about reality and should be treated as such—concepts act as commentary on reality rather than as reality itself. For the purposes of applied ontology-building in the natural sciences, such divergent definitions of concepts can lead to confusion. The realism-based approach, on the other hand, states that applied ontologies that are developed for scientific purposes should be built in such a way that their terms refer to universals in reality. The realism-based approach to ontology-building against which philosophers such as Cocchiarella and Merrill have positioned themselves has been formulated and extensively laid out by Smith and Cuesters (Smith, 2004b, 2006; Ceusters & Smith, 2007) and their resolution to the controversy is somewhat underwhelming. In their response to conceptualist attacks led by philosophers such as Merrill, Smith and Ceusters (2010) state that "Merrill's critique is of little relevance to the success of our realist project, since it not only reveals no actual errors in our work but also criticizes views on universals that we do not in fact hold" (p. 139). There is a sense in which conceptualism might be legitimate with respect to the status of language as it is expressed within consciousness however conceptualism offers no substantial criticism of the realist applied ontology project.

The thrust of the realism-based approach to ontology-building – as summarized by Merrill (2010), its most vocal critic – can be formulated as follows. First, realist positions believe in the view that specific terms are aligned not with concepts but with entities in reality (Smith, 2006). This sort of thinking is grounded in linguistics and can be aligned with the work of Katz (1966; 1981). Second, a realism-based approach to ontology-building should be able to plot nodes and edges on an ontology graph that corresponds not to concepts but to entities in reality (Ceusters & Smith, 2006a). Third, universals form the subject of scientific research; concepts form the subject of ideas, thoughts, and meanings. Universals are connected to instances in reality and are identified by discovering families of instances that share common properties (Ceusters, 2006). Lastly, for the realist, a kind is a part of reality that corresponds to universals or patterns in reality (Smith, 2006). The idea is that it is universals that provide for descriptions of many different particulars and that make science possible. Universals are what make terminologies possible.

While the nominalist position should be untenable in applied ontology work, the conceptualist position engenders some sympathy in the realist. Surely, concepts are real. Yet, they are not sufficient in describing the types of entities, attributes, and relations that should populate an applied ontology in the natural sciences, and should not the foundation upon which applied ontologies are built.

2.4 Realism-Based Applied Ontology is a Method

Some potential criticisms of the realism-based approach to ontology-building can be averted as long as the approach is understood as a method. The realism-based methodology “is based on the idea that the most effective way to ensure mutual consistency of ontologies over time and to ensure that ontologies are maintained in such a way as to keep pace with advances in empirical research is to view ontologies as representations of the reality that is described by science. This is the fundamental principle of ontological realism" (Smith & Ceusters, 2010, p. 139). Ontological realism is a method and not a philosophical doctrine; it borrows terms like universals and particulars from philosophy but "it does not stand or fall according to whether universals or types do or do not exist in some metaphysical sense" (Smith & Ceusters, 2010, p. 141).

Philosophers enjoy asking absolute questions that often create an absolute mess (Floridi, 2011, p. 74). The methodology of realism is meant to provide a productive solution to applied ontology work while utilizing tenets from philosophy, but it does not attempt to make a grand metaphysical claim. Rather, it uses the language of philosophy to show how to do good empirical work in the sciences relating to categorization. Building applied ontologies that last and that are successful is extremely difficult work. Their success is not found in solving longstanding realism vs. antirealism debates in philosophy but in the longevity and success of their application. A kind of pragmatism is practiced in the method of ontological realism, which seeks to provide the clearest and most widely-used standards for data integration and semantic interoperability.

2.5 Aristotle's Ontological Square

Philosophers have only recently engaged in applied ontological work in the computer and information sciences however ontology as a philosophical category has existed since Aristotle—arguing against Plato's theory of Forms—made a distinction between universals and particulars. The notions of universals and particulars are largely considered to be the foundation upon which much contemporary ontology-building is constructed. One famous scientist I spoke to illustrated that "the science of ontology is very old" and that it "started with Aristotle." Aristotle can arguably be said to have constructed a realist theory of categories in that he focused on the common properties of entities and how these might be grouped together; in Aristotle's ontology one finds that intelligible universals can extend across various domains. While Aristotle's ontology has been subject to debate among philosophers, many of the scientists involved in applied ontology with which I spoke recognize Aristotle's contribution to ontology research. Aristotle is often referenced in the scientific literature as the first "ontologist" and arguments have been put forward by scientists in favor of the view that Aristotle's philosophy can be seen as the first upper-level ontology (Schulze-Kremer, 2002) given that Aristotle's work in his Categories dealt with general first-order questions relating to items like substance, quantity, quality, relation, place, time, situation, condition, action, and affection; on Aristotle's theory, one can inquire into these basic categories to understand reality. Such categoricalism – the view that all of nature is categorical – suits ontology engineering given that such categories may adhere to multiple entities, attributes, and relations. Since the goal of ontology engineering is semantic interoperability between data, Aristotle's categories provide a good foundation for finding similarities between the seemingly disparate items from which data are generated and there is a link to rhetoric here in its contemporary form and the productivity of scientific rhetoric (Gooch, 1975; Depew & Lyne, 2013). Categories provide conceptual clarity when engaged in the actual practice of applied ontology-building so that items and definitions in the ontology remain clean and uniform.

The applied ontological interpretation of Aristotelian metaphysics holds that universals are inherent to their particulars and that the understanding of what is real is a matter of understanding the relationship between particular substances and their universal qualities. The universe is viewed as being comprised of types, universals, and categories that are hierarchically organized. According to Smith (2009), the most important universals in Aristotle's ontology are substance universals which pertain to what a thing is at all times at which it exists; however there are also accidents which pertain to how a thing is at some time at which it exists. These universals and particulars coupled with substances and accidents produce an ontological square (Figure 2.1).

	Substantial	Accidental
Universal	Second substance *man* *cat* *ox*	Second accident *headache* *sun-tan* *dread*
Particular	First substance *this man* *this cat* *this ox*	First accident *this headache* *this sun-tan* *this dread*

Figure 2.1 Aristotle's Ontological Square (Smith, 2009)

Such an ontological square captures Aristotle's early contribution to ontology and today many ontologists depend on Aristotle's categories, universals, and particulars to engage in applied ontology. Looking at the ontological square, for the realist the universal cat does not exist separately from the particular cat (such would be a return to Plato's theory of Forms). Rather, the universal cat is inherent to particular cats that exist. Similarly, the accident headache is a universal that adheres to particular headaches.

Aristotle’s ontological square endorses a realist theory of categories by virtue of the fact that, on the present ontological interpretation, a universal cannot exist that does not already inhere in some particular substance or accident. Once it is partnered with such a realist approach to universals and particulars, categoricalism no longer appears to be as abstract as it may seem and difficult philosophical questions pertaining to the reality of concepts become less relevant. The conceptualist position might hold that a universal originates in the thought of the epistemic subject; however such would not be a realist theory of categories. Universals do not originate outside of their particulars and as such the Aristotelian metaphysics endorsed by ontology engineers should suggest that the content of ontologies should be realism-based rather than concept-based entities, attributes, and relations—yet many ontologists continue to embrace the term concepts in the technical literature. It should be acknowledged that Aristotle did not think of his categories in terms of information or realism yet this should not prevent ontologists from retroactively reading his work as realism-based and informational in nature.

To emphasize the relationship between universals and particulars, one leading ontologist recommend to think of the link between universals and particulars in terms of a principle of instantiation. The principle of instantiation is used to make sure that users of the realism-based approach to applied ontology-building see types, categories, and universals not as abstract entities that exist in a magical realm outside of reality that exists beyond contact from empirical observation, but instead as firmly planted in the world and aligned with those entities with which scientific data is connected. A conceptualist might then inquire into the nature of hypothetical scientific objects that have not been proven by science and how they should be included in an ontology using the realism-based method. For example, should such entities wait for science or rely on concepts? The answer provided by such realists such as Smith and Ceusters (2010) is rather straightforward, and rests "on the recognition that language can clearly still be used to communicate – in some sense – even where putative referring expressions fail in their reference. Some people assert their beliefs in the existence of unicorns. All such beliefs are false. But the beliefs exist just as do other beliefs; they can be communicated; and they can also be represented" (p. 179). Beliefs in a non-existent entity can be realistically represented in an ontology even when that entity itself cannot be represented in the ontology. The significance of this is that a space can be provided for the inclusion of non-existent or hypothetical entities in ontological realism so long as they are referred to as beliefs, a practice which eliminates categorization errors in taxonomy.

2.6 Husserl's Formal Ontology – Formal Logic Distinction

Formal ontology is used by philosophers and computer and information scientists to refer to ontology as a discipline. The term is included, for example, in FOIS and philosophers have used it ever since Husserl began referring to a formal ontology at the begging of the twentieth century. The applied ontologists I spoke to recognize the work of Husserl and describe him as the inventor of formal ontology as a discipline distinct from formal logic. What Husserl accomplished with this was a double distinction in the form of separating descriptions of material things that exist in the world from thinking about the abstract structure of those things and the separation of formal logic from those abstract ontological structures. Husserl famously advocated for a 'return to things themselves,' which is often interpreted to mean a return to thinking about individual objects as they exist in the world and can be apprehended within phenomenological reality. As such, formal ontology began as a field connected to perception and common reality. One well-known ontologist said "Husserl was the first philosopher to use the phrase 'Formal Ontology' in his Logical Investigations, which is also important for my development. I was very much influenced by Husserl." Husserl showed how philosophy and science had become detached from the life world of ordinary experience and wrote at length about topics related to Aristotle's categories, including subjects such as universals, particulars, and meaning, yet he focused on them from an explicitly ontological perspective, seeking instead to provide a universal account of ontological structures. Husserl's *Logical Investigations* (1900/2001; 1901/2001) provided an early account of his ontology that would be developed throughout his later works, particularly his *Formal and Transcendental Logic* (1929/1969); in it Husserl offers a clear distinction of a formal discipline of ontology that should be viewed independently from formal logic. The formal status of Husserl's ontology is meant to imply that the ontology can extend to all entities, attributes, and relations without being impeded by specific knowledge in various domains of application.

Husserl saw logic as concerned with meaning and the deductively closed collection of meanings that constitute scientific theories. He takes this a step further by arguing that it is only until we have a complete theory of objects to which meanings refer that we can have a complete scientific theory; thus the unity of logic depends on a connection of meaning and a connection of objects to which meanings are directed (Smith & Smith, 1995a; 1995b). Formal logic is related in the first instance to meaning categories such as subject, predicate, proposition and any terms that are veridical in nature; formal ontology concerns object categories such as entity, attribute, and relations, including such items as parts, wholes, and states. Logic can be said to provide the concepts that belong to a unified theory between meaning and object; logic's truths are necessary truths related to such categories. Formal ontology provides models that are similar to formal logic in that they form complex structures in non-arbitrary, rule-governed ways. Such a process is independent of domain specific knowledge and thus can be applied in any number of cases.

The sociologist Robert Poli has worked extensively on ontology for knowledge organization (1996) in both the applied and philosophical sense, as well as on basic problems relating to theories of levels of reality (2002). He has also provided one of the most thorough readings of Husserlian ontology, emphasizing that Husserl's ontology is a formal system that sits next to formal logic. It is a rule-governed system that is supposed to correspond to the material entities that are the referents of concepts. This interpretation of Husserl, which is by far the most common and accepted interpretation of his ontology, still posits that concepts are to be found behind formal ontology and ontological commitments (Figure 2.2). Formal ontology has different interpretations throughout history and in some cases the phrase is used by analytic philosophers when referring to formal logic, particularly with regard to scientific theories and structures. Husserl's formal ontology is separate from formal logic and is practiced according to this distinction by computer and information scientists interested in ontology work (Poli, Healy, & Kameas, 2010).

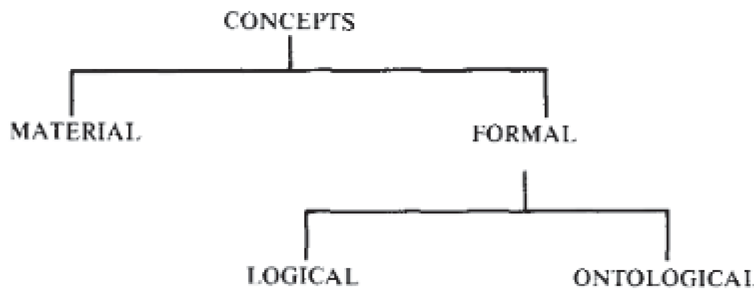


Figure 2.2 Poli's Logical-Formal and Ontological-Formal Concepts (Poli, 1993)

Such a conception of formal ontology made its way from Husserl to modern information science and the beginning of the applied ontology revolution. Guarino and Poli (1995) note that the "International Workshop on Formal Ontology in Conceptual Analysis and Knowledge Representation" held in Padova, Italy on March 17-19 1993, was probably the first interdisciplinary initiative aiming to "explore the connections between philosophers belonging to the tradition of Husserl […] and people working on principles of knowledge representation and engineering" (p. 623). The proceedings show that the connection between Husserlian ontology and computational knowledge representation "proved to be very influential to the communities involved" (p. 623). Computer scientists saw the value in viewing the distinction between formal logic and formal ontology for showing that ontology should deal strictly with the interconnections of things, with objects, properties, parts, wholes, relations and collectives, while formal logic should deal with interconnections of truths, with consistency, validity, and statements like and, or, and not. Husserlian formal ontology views ontology as something that should exist in every domain of reality, and this is what makes it formal. It is not material in the way that physics and biology are in that it does not deal with material entities but rather with ontological things in the abstract philosophical sense. Formal ontology deals with formal ontological structures while formal logic deals with formal logical structures. The word formal merely means that formal logic and formal ontology are meant to obtain in all spheres of reality. Here, material should be understood as regional structures within specific domains; formal ontology is domain-neutral and concerns parts of reality that are shared by all material domains. Material ontology deals with features which are specific to given domains, things like brains, buildings, etc.

As Poli's logical-formal and ontological-formal concepts show, concepts remain embedded at the top of the hierarchy, and in this there also remains a sense in which the origin of concepts stems from the mind of the epistemic subject who apprehends material, or formal-logical/formal-ontological structures. As is well-known by philosophers, Husserl is the godfather of phenomenology, and his phenomenology is centered on the epistemic subject's intentional relation to objects; for Husserl, it is the nature of thought to be intended for an external object. On this view, different concepts can apprehend objects in different ways. This aspect points to the descriptive value of Husserlian phenomenology for the material world, but in the realm of applied ontology it is Husserl's separation of formal-logic from formal-ontology that has left a lasting contribution to the way ontology is studied today. However, the positing of concepts at the top of the ontological pyramid remains problematic for the modern-day practice of applied ontology-building and the realist project.

2.7 Ingarden's Ontological Essence of Time

The Polish philosopher Ingarden was a student of Husserl who viewed ontology as a science of the possible ways of existence. When I asked one ontologist which books or philosophers they considered to be important to ontology, they stated that "Probably the philosopher whose work has influenced the ontologies which I build most centrally is Roman Ingarden…He would be the person whose views come closest to my own." When this ontologist was studying at Oxford they had specialized in analytic philosophy and eventually became interested in looking to alternatives; this is how he discovered the work of Ingarden, which is to say more or less by accident one day in a library. The ontologist became interested in Ingarden because of his work on aesthetics but discovered that Ingarden wrote a rather large book on ontology, parts of which were translated into English in 1964. When the ontologist discovered that book and saw what Ingarden was doing, they told me, "I realized that I was an ontologist." They said it happened within seconds, that they had the book in their hand, opened several pages, and realized that applied ontology was what they wanted to do. "And I've been doing it ever since," they told me. "Which means now something like forty-five years."

Ingarden is an interesting figure; a philosopher who many other philosophers have for a large part ignored for one reason or another, his philosophy is the subject of admiration for quite a few scientists in applied-ontology work. Ingarden's monumental *The Controversy over the Existence of the World* (1947/1948), originally published in Polish, is a masterwork in ontology that sought to define the world (and the possible) in terms of ontological categories. The edition that our Ingarden-influenced ontologist found was an English translation of only a portion of *The Controversy* (which is in two volumes) entitled *Time and Modes of Being* (1964). The “controversy” in the title of Ingarden’s book refers to the one that has been generated in the debate between realism and idealism. Importantly, Ingarden introduced the notion of an objective, realism-based time into formal ontology and this seems to have been his key contribution and the reason for his profound influence. Unlike Husserl, Ingarden moved beyond concepts and phenomenology to create a realism-based formal ontology that included time as an objective universal category. Building on Husserl’s formal ontology and in disagreement with Kant, Ingarden writes that

> The question at issue here does not belong to a general theory of time or to a general theory of existence, but is rather a problem which in this context – where we examine the ontological problem of the existence of the real world – is of vital interest to us. For – true or not! – the real world as we grasp it in pre-philosophical, everyday experience appears to be organized in such a peculiar fashion that anything and everything that occurs within its unity is somehow temporal, or is at least bound up with time. But even if that should turn out to be a transcendental illusion-say, in Kant's sense-still, the problem of time cannot be left out (2013, p. 227).

Ingarden claims that according to his realism-based ontology, entities can be divided into the temporal and atemporal, and he spends the majority of Time and Modes of Being brilliantly combing through the differences between objects existing in time (things in particular), events, and processes. It is interesting to note that the ontologist I spoke with was surprised in retrospect by how similar Ingarden’s continental philosophy was to that which was being studied by analytic philosophers in England at the time, however Ingarden’s emphasis remained focused on logically partitioning reality rather than on the abstract language games that were often practiced by English philosophers at the time. Ingarden, in a footnote in The Controversy, remarks that he was unaware of the work of analytic philosophers such as Alfred North Whitehead (1861-1947) while writing the book.

Ingarden states that the temporal nature of the mode of being of an event consists of the occurrence or coming-into-being of some state of affairs or of some object-involving situation (1947/2013, p. 229). He provides as examples of events the collision of two bodies, the arrival of a train at a station, a lamp's lighting-up, a person’s death, and the like. The word event is also used in common language when referring to things such as a battle or other historic events. However, Ingarden argues that those types of cases are actually processes of relatively brief duration which display an inner unity (a coherence of phases) and are contrasted with longer-lasting processes.

On the temporal nature of the mode of being of processes, Ingarden writes that examples would include things like a specific, concrete movement of a material mass in space such as a runner’s dash at a track meeting, the evolution of an organism, the life of a human being, and any and all activities and transactions of a purely physical (as well as psychophysical) nature. In every process (e.g. a well-defined movement) one should here distinguish between the continually growing totality of phases and the object constituted in them in the course of time as the processes’ peculiar subject of properties (1947/2013, p. 235). However, both categories make up the single entity in which they are distinguishable only as two different aspects. The general constitutive property of this process-object is that the growing totality of phases unfolds in time. This means (1) from phase to phase – from inception to conclusion – the process runs its course in ever new time intervals, and (2) the totality of the process-phases grows constantly until its closure, and – in contrast to the event – it cannot in virtue of its essence be contained in a single instant, in a single present.

Finally, the temporal nature of the mode of being of objects involves some arbitrary thing; a stone, a house, or a mountain can each be taken as example of this sort of object. Living beings such as cells or trees, animals, humans, as well as specific humans such as President Barack Obama can all be considered objects. Living beings (especially the cellular) pose certain difficulties if they are to be sharply contrasted with other temporally determined objects. But on closer inspection, Ingarden reveals that it is precisely they that enable us to discern the radical distinctiveness of persistent objects from events and processes; persistent objects differ from events by outlasting the individual instants in which events are confined (1947/2013, p. 251-252). This also applies to processes. Meanwhile, it is precisely in how a persistent object out-lasts individual instants that its deep disparity from processes is exhibited; a process does it in such a way that its currently active phase passes over into a wholly new one—though essentially inseparable from it; the former prolongs itself continuously into the latter. In contrast, a persistent object remains as identically the same in the incessantly new instants of time for as long as it exists. If one were to find something new in it in the newly incipient instants of time, it is either processes that are existentially interconnected with it, which sometimes play out internally, or certain events that take place in the object. This is to say that both can elicit new properties in it, or entire ensembles of them. But the persistent object itself, which serves as existential basis for the various sorts of entities that frequently coexist with it, remains, so to speak, as the same old thing that already existed earlier, in the previous, elapsed instants.

Ingarden's analytically rigorous yet realism-based approach to objective time implies that objects, events, and processes exist at various levels of granularity, and his ontological project is to formally define such categories to assist science. Ingarden writes that "it appears to be likely that the real world, or at least what exists in it, is temporally determined. At any rate, time-determination belongs among the primal attributes of entities existing in the real world, much as it may be no more than a "transcendental illusion" in Kant's sense" (1947/2013, p. 280). Taking a stance against Kant, Ingarden's ontology permits of an objective universal conception of time that, once combined with Husserl's formal ontological method, expanded the class of entities that could be permitted in an ontology. It is this characteristic of Ingarden's work that seems to have influenced the ontologists I spoke to. In the turn from idealism to realism, Ingarden rejected the notion that the essence of things like time could exist only as an allusion in the mind in favor of viewing time as a real, constitutive parts of the universe.

2.8 Ontological Realism in Modern Applied Ontology

The work of philosophers such as Aristotle, Husserl, and Ingarden can be said to rest upon the realist assumption that a single consistent ontological theory can comprehend reality at a multiplicity of different levels of granularity. There are many other philosophers that are relevant to a discussion of the realist ontological project (Simondon, 2005; Floridi, 2011) but covering them would be beyond the scope of this dissertation. However, the work of Quine seems to be a real turning-point in ontology work for some of the computer scientists I spoke to. Quine, some philosophers explained to me, famously did not believe in the existence of properties. His ontology really supported only a realism about nouns and his theory of ontological commitment sought to study not what there is but what sciences believe there is when logically formalized. Quine thought that the way to understand the ontological commitments of a scientific theory is to translate the scientific theory into predicate logic form; this could then allow one to read off the ontological commitments by looking at all the predicates needed to capture the existential assertions of the science. In Quine's terms, to be in the ontological commitment sense is to be the value of a variable in a scientific theory formulated using predicate logic.

Quine and many of his contemporaries in the analytic tradition of philosophy established a common understanding according to which the use of first-order logic as a technique of philosophy should accompany the acceptance of a nominalist view as concerns the variety of things to which constituent terms in first-order logic are allowed to refer. The view that came to be adopted in much philosophy concerning ontology found that all terms in first-order logic should refer to individual objects (atoms, cells, persons, etc.) with the outcome of a restriction on the available expressive resources of first-order logic. According to Smith and Ceusters (2010), universals fall outside of the domain of what can be referred to within the framework of first-order logic.

> Because terms in received FOL [first-order logic] range exclusively over individual objects such as molecules or cells or people, such terms cannot be used to refer to universals, or to anything general or repeatable. And the predicates in FOL cannot be used to refer to such entities either – because they cannot be used to refer to anything at all (p. 158).

The realism-based approach to ontology construction depends not only on first-order logic but also on what Smith has called first-order logic with universal terms. First-order logic with universal terms is different from first-order logic in that it expands the types of entities to which first-order logic can refer while also majorly restricting the group of allowed predicates, abolishing all predicates of the typical sort (*is_a* man, *is_a* thyroid, etc.) and admits rather only a small amount of formal predicates, including two-place predicates (relational). First-order logic with universal terms allows terms to refer to independent and dependent continuant particulars and occupant particulars and also universals in each of these categories (Smith, 2005).

A summary of the realism-based methodology is found in Smith and Ceusters (2010). Therein, relational predicates on the level of instances are described, as well as relational predicates on the level of universals or types, and bridging universals and particulars (Figure 2.3).

Part_of(x, y), for: individual x is part of individual y
Member_of(x, y), for individual x is a member of individual collection y
Inheres(x, y), for: individual x inheres in individual y
Precedes(x, y), for: individual process x precedes individual process y
Has_Participant(x, y), for: individual thing y participates in individual occurrent x
Has_Agent(x, y), for: individual thing y is agent of individual occurrent x
Realizes(x, y), for: individual process x realizes individual function y.

is_a(x, y), for: every instance of universal x is part of some instance of universal y
part_of(x, y), for: every instance of universal x is part of some instance of universal y

Inst(x, y), for: individual x instantiates universal y
Extension(x, y), for: individual collection x is the extension of the universal y

Figure 2.3 First-Order Logic with Universal Terms (Smith & Ceusters, 2010)

Temporally indexed versions of each are included in the ontology. The outcome of expanding the scope of permitted referents in first-order logic to include universals is that it allows for the simulating some of the expressive possibilities of second-order logic within the framework of first-order logic, including relations.

The realism based approach to applied-ontology building is a radically new approach. Previously, concept-based ontologies did not include universals and as such many of them fell into disarray. The rules included in first order logic with universal terms combined with the LoA approach allow for greater semantic interoperability to produce a realism-based perspective of applied ontologies that can then be used to combine different types of scientific data for the purposes of progressing science.

Understanding the technical theories that inform applied computational ontologies opens the black box of ontology up to see its inner logic and rational, making clear the mechanics behind ontological decision-making and development. However, the rhetorical and practical work of ontologists themselves contributes to work in the community development of ontology knowledge-sharing. Gieryn (1983) has emphasized that scientific communities are bound by rhetoric and practice that influence scientific methods and research. Chapters three and four address this in the context of communities of practice who are engaged in applying the method of ontological realism to science and basic research. Chapter three offers a brief history of contemporary communities involved in applied computational ontology research while chapter four offers a specific look into communities involved in the development of BFO.

CHAPTER 3. SCIENTIFIC ONTOLOGY AS COMMUNICATION

3.1 Chapter Summary

This chapter provides a theory of community practice, describes individuals, groups, and institutions involved in ontology work at the local and global levels, and argues that ontology should be studied as a form of community practice. Communities of practice have been described in Science and Technology Studies literature as the formal and informal connections that emerge when interdisciplinary groups of knowledge seekers form around an emerging technology (Wenger, 1998; Schiavone, 2014). The chapter describes various groups and organizations involved in applied ontology research and provides some of the history behind their origins and development. It describes some of the communities that have been active since the 1990s who formed around and participated in applied ontology research, particularly those who have embraced the realism-based method, including the conferences and international organizations that have come to embrace the realist method.

3.2 Ontology as a Community Practice

Communities of practice in the fields of computer and information science have been described in the Science and Technology Studies literature as consisting of four types (Cox, 2005), including the socialization of new group members into a knowledge field by a form of apprenticeship (Lave & Wenger, 1991), creating knowledge in interdisciplinary groups that form in resistance to old hierarchies of power in institutions (Brown & Duguid, 1991), informal connections among individuals that emerge through mutual engagement on knowledge ventures (Wenger, 1998), and informal horizontal management groups across institutional boundaries (Wenger, McDermott, & Snyder, 2002). In the relatively short history of applied computational ontology work, communities of practice seemed to have emerged informally among researchers through their interests in ontology, thus the story of ontology fits into Wenger's (1998) informal account.

In many ways, the practice of applied ontology building is a black art. How does one go about organizing scientific knowledge if there are no theoretical foundations or practical principles? To come up with such foundations and principles, ontologists work with computer scientists, philosophers, communication specialists, engineers, logicians, and many others. To work in their trade requires that ontologists get along with a somewhat mysterious blend of specialists. Merrill (2010, p.105) accuses Smith and Ceusters' realism-based approach of being "neither science no philosophy." In their response (2010), Smith and Ceusters write that in suggesting such Merrill "hits the nail exactly on the head." In propagating the realist methodology they claim that they are "indeed engaging in a novel interdisciplinary activity that involves elements of both of these, and also of computer science, politics, community organizing, sociology, logic, and other black arts." Smith and Ceusters describe coordinated ontology work across a large scale as being so difficult that they are "happy to draw on any means that will help us to achieve our ends." Building on such literature in the realism-based approach to ontology construction, I argue that ontologies are commonsensical and real but also that they require significant communicative effort to be maintained. The fact that ontologies require communal maintenance is something that has been overlooked by researchers in communication and media studies. "I increasingly see myself as being a practitioner, or a scientist, working in the discipline of ontology," one scientist told me. Such individuals and communities centered on ontology work should be studied by communication researchers.

I asked one scientist if it is correct to say that ontology is trying to solve a communication problem amongst scientists. "Yes. That's true," they said (J. Doe, phone interview, September 9, 2015). Many realism-oriented ontologists explained to me that the ToB problem in ontology is going to be a problem for multiple fields and disciplines in the future—from statistics to biology, and from physics to farming. The more computers are introduced to the world and the more computer-driven daily activities and research activities become, the more ToB problems are going to have to be addressed. Somehow, all of this data will need to be interfaced in such a way that they become semantically interoperable. Problems will have to be addressed first not in those areas which have Big Data, but in those areas which have heterogeneous big data. There is less ontology development in physics, for instance, than there is in biology. Physics has huge amounts of data in the field of astronomy (there is work in the ontology of astronomy) but it is not as important as work in the ontology of biology. The data of biology and the data of medicine is so fantastically heterogeneous that researchers are forced to try and find ways of making it comparable across species, diseases, experiments, and across chemistry. Chemistry is an area where we now have good ontological resources but applied ontology work is still in its infancy in many domains.

There is some agreement on the notion that ontologies can be treated as collectively agreed upon scientific theories, rather than strictly as engineering artefacts (Eschenbach & Grüninger, 2008b, p. v). Up to fairly recently, the term *controlled vocabulary* was used in much of the scientific literature to describe the practice of organizing words and phrasing to better enable the indexing and retrieval of content through computerized searching. Similar to ontology, controlled vocabularies usually included different terms and domain scopes. Unlike ontologies, an emphasis on community was not typically associated with a controlled vocabulary. Today, the term ontology has recently taken over from controlled vocabularies (Bodenreider, 2008) in the domain of biomedical informatics (Figure 3.1).

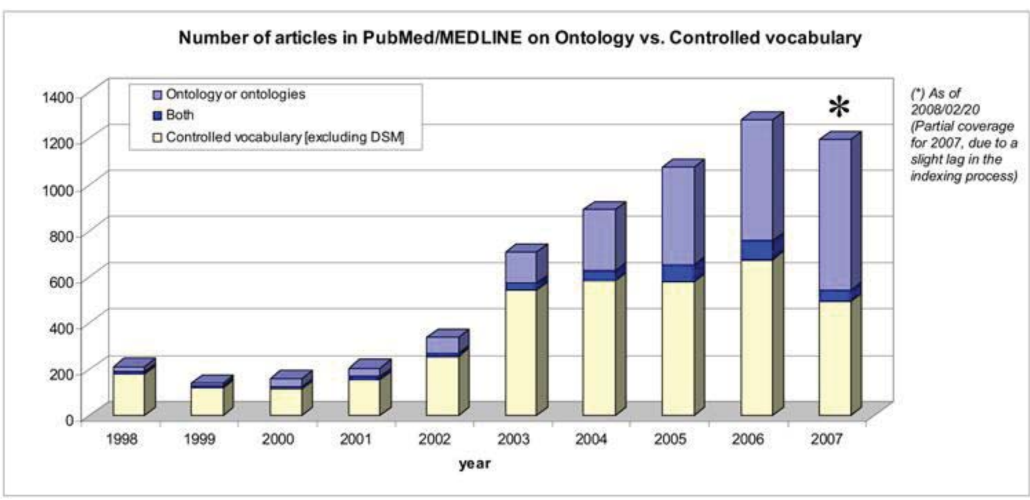


Figure 3.1 Ontology vs. Controlled Vocabulary in PubMed/MEDLINE (Bodenreider, 2008)

Even when terms such as *knowledge engineering* were still being regularly used in the 1980s and 1990s, Uschold and Gruninger (1996) noted that the construction and use of ontologies involved different backgrounds, languages, tools, and techniques that often acted as a barrier to effective communication among people, organizations, and systems. They noted a need for the development and implementation of techniques to enable shared understanding and communication between wildly divergent teams of researchers.

To produce more and better ontology work, the ontology-building process must be seen as a communal endeavor. Yim (2015) argues in favor of the communicative approach to applied ontology, suggesting that ontologists should diversify the membership of their community of practice, expand penetration into education, keep intellectual property rights open, and continue bootstrapping. The value of the communicative approach to ontology work is found in the interdisciplinary nature of applied ontology-making. Ontologies require not only philosophers to work on the logic of the ontologies but also communication specialists (ontologies need to be advertised, disseminated, measured, reviewed, etc.), computer and information engineers (ontologies need to be instantiated using computer science and programming), managers (ontology is a team science involving different research groups), and many others.

3.3 Realism-Based Ontology Groups and their Institutions

Institutions also play a large role in that realism-based applied ontology work is a costly and time-consuming process that requires significant resources and infrastructure. Today, the realism-based approach to applied ontology-building is practiced within a variety of institutions in the northeastern United States, with the majority of them located in Buffalo. As one researcher there put it to me, Buffalo is quickly becoming “the Silicon Valley of ontology research,” with much of the ontology work attached in some way to the University at Buffalo. “I think now most people would say it is Buffalo which is the most important player,” another scientist told me. One ontologist also stated that “Buffalo is probably the largest single community of people who are building, or maintaining, or using ontologies.” Part of the reason for this is that ontology work requires large groups of teams and institutions working in close contact. Buffalo has been or is currently home to many ontology research groups, centers, and individuals, including the Ontology Research Group (ORG) (directed by Ceusters, Smith, and Louis Goldberg), the National Center for Ontological Research (NCOR), and the Buffalo Center for Ontological Research (BCOR).

The ORG is a part of the New York State Center of Excellence in Bioinformatics & Life Sciences (CBLS), which is housed in the University at Buffalo. Researchers at the ORG have experience in ontology research in many different disciplines, including the review and editing of applied ontologies in biomedical informatics as well as skills in referent tracking and language processing. Work within the ORG is divided into three specialized units: The Ontology, Logic and Technology Unit (OLT) conducts primary ontology research and development in biomedicine; the Referent Tracking Unit (RTU) specializes in software research focusing on electronic health records; and the Qualitative Spatiotemporal Reasoning Unit (QSR) applies ontological methods that come from qualitative reasoning in fields such as Geographic Information Systems to improve upon representation systems and other forms of image data. The goal of the ORG is to help scientific researchers working primarily in biomedicine by providing single databases in specific domains that can be computationally processed. The ORG’s mission is to create high-quality domain ontologies that can facilitate translational research.

The NCOR was created in 2005 with the aim of conducting ontological research by building tools and measures for ontology evaluation and quality assurance. The National Institutes of Health thought that it was important to support the Gene Ontology (a specific domain ontology) and similar efforts because of the importance of the Human Genome Project for medical research. The principle investigator of the NCOR was Mark Musen (Musen is currently at Stanford) and Smith served as co-principle investigator. The pair thought that it would be important to have something similar for non-biomedical ontology and so they NCOR for areas which were at that stage restricted to a small amount of military work and a larger amount of theoretical work in domain and discipline neutral ontology work. Within a short time, the NCOR moved to be fully located in Buffalo and Smith became the director. Currently, the NCOR does a lot of work for military projects and has done most of its work within the military ontology research area. However it is not restricted to military projects and is still active in areas such as financial services, economics more generally, and legal ontology.

Ontologists working at the University at Buffalo participate in the center and its activities and collaborate with scientific, private, and public institutions in the USA and around the world by organizing ontology-themed research activities, conferences, publications, and funding opportunities. NCOR operates as an infrastructural hub that enables the coordination and review of organizations that use ontologies in many different fields, including national defense and intelligence, management, and healthcare. The center provides resources for those engaged in applied ontology by helping them find funding and establish interdisciplinary teams. It also provides consultants for ontology-related projects, particularly in security and healthcare. They engage in training and outreach that are designed to spread ontology research around the world and to include more institutions and individuals in the applied ontology-building process.

The BCOR, also housed within the University at Buffalo, contains many different faculty projects that span across a variety of different departments. BCOR specializes in collaborative work and has secured major funding from the National Science Foundation, the National Institute of health, North American defense agencies, and the European Union. Goals of the BCOR include providing a forum for philosophers and applied ontologists to work together on multidisciplinary research projects. The BCOR focuses primarily on biomedical ontology—molecular biology and biochemistry, functional genomics, and proteomics, specifically.

In Europe, the Institute for Formal Ontology and Medical Information Science (IFOMIS) at the Philosophy Institute of Saarland University is perhaps the largest organization specializing in ontology. Its primary research activities are the investigation of the basic philosophical tenets of formal ontology, the development of specific domain ontologies, and the development of data integration techniques for ensuring semantic interoperability. IFOMIS is home to multidisciplinary groups of researchers including members from philosophy, computer and information science, and medicine.

IFOMIS was founded in 2002 after Smith won the 2001 Wolfgang Paul Award of the Alexander von Humboldt Foundation (with additional funding supplied by the Volkswagen Foundation and the European Commission). The Wolfgang Paul Award was granted then for the first and last time to fourteen promising researchers. The award was worth four and a half million German marks (the equivalent of just over two million Euros) and was the most valuable research award for a scholar at that time. It provided winners with the freedom to pursue their research projects away from administrative constraints at a German academic research institution and to create their own interdisciplinary working groups of researchers. The award was financed by the Federal Ministry of Education and Research and were made available from a federal government program that existed at the time named the Future Investment Programme. In total, fourteen scholars received the award (a total reaching almost twenty-six million Euros), and Smith was one of only two recipients from the humanities, having been trained as a philosopher. The other recipient from the humanities was a linguistics professor. The majority of the awards went to individuals in the fields of physics, mathematics, engineering and geoscience.

At the beginning, IFOMIS was established at the Faculty of medicine in the University of Leipzig but it eventually relocated in 2004 to Saarland University in Saarbrücken to capitalize on interdisciplinary collaboration at the frontier of computer and information science research. IFOMIS set the task for itself of advancing research in ontology in the field of bioinformatics and to prove how the knowledge of philosophers can be greatly beneficial for such fields. Doctoral researchers in various fields including philosophy but also medicine, linguistics, and computer science have interacted with IFOMIS and its training and research modules.

IFOMIS is largely responsible for bringing about a worldwide transformation in the logical development of ontology research and have developed highly refined techniques for evidence-based applied ontology development which have spread in popularity and are now used worldwide by various well-known and highly regarded ontology groups. The methods developed at IFOMIS led to the development of the BFO, a top level ontology that organizes multiple domain ontologies and which is now serving as an integrating framework for a large variety of ontology projects. In 2005, Smith joined the Gene Ontology Consortium to found the Open Biomedical Ontologies library, which later in that same year become institutionalized as the OBO Foundry. Smith is still heavily involved in OBO Foundry activities and in the development of BFO principles. The OBO Foundry and the BFO will be discussed in the second half of this dissertation.

3.4 The Data Smiths

I asked one ontologist what they viewed as being some of the biggest challenges facing ontology development today. They said that there were three main issues. One issue is that the ontology software is becoming more usable, which means that a lot of people are using ontology software to build ontologies who do not know how to build an ontology. There is an influx of bad ontology content, which is giving the people who have been building good ontologies a bad name. Another problem is that there is a shortage of trained ontologists with the right kinds of skills. The ontologist told me that any student of theirs who was trained in ontology immediately gets a job because there is such a need and that this will continue to grow over time. The other issue is that ontologies are becoming rather standard now. Lots of people see the need for an ontology, including institutions such as the NIH and other funding agencies. The problem is that institutions like the NIH do not fun ontology research as they once did because they view ontology building as a kind of standard, as being infrastructure. Primarily, they are concerned with and want to fund original research. They do not want to fund infrastructure. This, there is a problem in getting the right kind of funding to enable ontology development work to take place. This is a problem that applies not just to ontologies, but to maintaining large database. If Big Data is going to continue to grow and be a prominent issue in research, the ontologist explained, then it is going to become harder and harder to keep Big Data in ways which allow it to be used properly because as it grows and costs more to keep it. Unfortunately, many funding agencies do not want to pay money to databases.

There are a handful of individuals who have been instrumental to spreading the realism-based method to applied ontology engineering—they are the data smiths who focus on the practice of increasing semantic interoperability among diverse data sets. Smith and Ceusters are two of the most important individuals working in applied ontology today; they are the creators and main proponents of the realism-based method to applied ontology research. Both work at the University at Buffalo and have at some point or another been associated with ORG, NCOR, BCOR, or IFOMIS.

Currently, Smith is SUNY Distinguished Professor of Philosophy and Julian Park Chair in the Department of Philosophy at University at Buffalo (Smith is also Research Director of IFOMIS). He studied at Oxford and received his doctorate from the University of Manchester and is currently also Director of the NCOR, as well as Adjunct Professor of Biomedical Informatics, Computer Science, and Neurology at Buffalo. Smith edits one of the oldest philosophy publications in the word—he is Editor of *The Monist: An International Quarterly Journal of General Philosophical Inquiry*. He is a well-known contributor to both philosophical and computational ontology and has authored over four hundred and fifty articles on ontology and ontology-related topics. His research has been funded by the United States' National Science Foundation (including those of Switzerland and Austria), the National Institutes of Health, and the European Union, among many others (including the Wolfgang Paul Award of the Alexander von Humboldt Foundation). Smith works primarily in the application of ontology in biomedicine and bioinformatics. He is also Coordinating Editor of the OBO Foundry.

The world's first institute of applied ontology was founded in Padua but then moved to Trento and is now called the Laboratory for Applied Ontology (LAO). Smith spent a semester working at LAO in 1993 and has been collaborating with the LAO team ever since (there is also a branch of the LAO in Rome). There is another laboratory for ontology, Labont, in Turin, with whom Smith has also collaborated. Smith's work with the Italians led him to discover the work of Hayes (Hayes & McCarthy, 1969). More generally, he discovered that there was a community of people in the AI world who were trying to solve a problem, which can be summarized as being a problem in robotics. How do you build a robot which would have the same common-sense understanding of the physical world that a human being has, and also a common-sense understanding of things like economics and ethics, and so forth? The robot is going to have to be able to do things like buy salad in a restaurant, and to do that requires a commonsense knowledge of economics, ethics, of politeness, salad physics, tomatoes, and so on. There was a big effort known as "formal theories of the commonsense world," in which Hayes was one of the most important figures, which was attempting to use formal logic to represent reality (Hayes & McCarthy, 1969). Since then, Smith himself has also produced work in this field (Smith 1995a, 1995b). This early work is one way of describing what ontology is trying to do. Smith started to write papers on naïve or formal theories of common sense. This work was partly theoretical and partly based on his historical work on people like

Husserl, who did in fact work very seriously on formal theories of common sense without having formal logic as a tool. That work led Smith to become interested in the ontology of geography, formal theories of space, common sense space, political space, and so forth. As the result of the work he did on the ontology of geography, Smith received the prize from the German government to found an institute on ontology research in Germany, and that is when the idea of working in biology and medicine became interesting to him; he had resources in Germany and wanted to do something ambitious.

Since Smith has been in Buffalo, now for more than 20 years, he has been working on ontology and has brought a number of collaborators and colleagues, including senior and junior faculty positions over the years. They have students and researchers that make up a large community of people doing biomedical ontology in Buffalo. There is also now a slowly growing community of people working for military ontology projects in Buffalo. Initially, the main institution around which Smith and his team built this collection of people was the CBLS. Now, there are several institutions, including the newly created Department of Biomedical Informatics, the head of which is himself somebody who is engaged in doing ontology work, amongst other things. There is also the Institute for Healthcare Informatics, which is engaged in a lot of ontology based in health and patient data research not only in Buffalo but in New York State and beyond. As a result of this growth in the numbers of institutions, Smith’s role is not tied to any one of these. He has an adjunct professorship in the Urology Department and in the Biomedical Informatics Department, and also a joint appointment with Computer Science. Smith works with researchers in other departments as well.

The goal of ontology is to make large heterogeneous bodies of data discoverable, comparable, and capable of being handled either by a computer or by a human being. The first discipline which really faced that issue and which had Big Data that they needed to grapple with was biomedicine. The Human Genome Project and the project to decipher animal genomes (the mouse genome, fly genome, and so forth) presented biologists and researchers doing clinical science with a difficult problem. The data they needed to use to do biology was gigantic sequences of chemical symbols with no obvious biological meaning. They needed to find ways of making genomic data meaningful, discoverable, and comparable. For example, this would allows them to compare the results of doing experiments on mouse diseases with the possible consequences of performing the same actions on human beings when it is illegal to do experiments on human beings with the same kind of freedom that is allowed when working with animal models. The animal model community, which was by the time Smith created IFOMIS already an informatics based community, conceived the first successful ontology in the modern sense, which was called the Gene Ontology. Smith organized a meeting on the Gene Ontology; he could see its importance. It was a successful artifact for researchers who wanted to do information driven biomedical research. Smith also could see that the Gene Ontology was, from a logical point of view, incredibly bad. The definitions were bad; they were circular or they were worse than circular. The researchers involved in the Gene Ontology did not understand the basics of first year logic when they created the ontology. They created something which was very successful, but which needed a logical structure. The leadership of the Gene Ontology visited Smith’s institute in Germany. The outcome of that meeting was that the ideas which he’d been working on for logically based ontologies became used by the Gene Ontology community as a way of increasing the logical structure of the Gene Ontology itself. Due to the influence of the Gene Ontology, Smith thereby became influential in the area of information driven bioinformatics and biomedical informatics more generally. Quite a significant amount of the work which has taken place in biomedical informatics since then in the field of ontology has been influenced, one way or another, by Smith.

When Smith won this prize from the German government in 2002, he got an email on the very next morning after the prize was announced from Ceusters. Smith did not know about him at the time and knew very little about the biomedical world. That was one of the reasons why he chose biomedical ontology as the focus for the institute that he was founding—it was a new area and gave him the opportunity to apply the work he had been doing in naive physics and geography to an area that was much more challenging and stimulating (for him). Ceusters was a leading figure in bioinformatics and medical informatics in Europe and had his own company focusing on the use of ontologies to support medical natural language processing, natural language translation, and natural language analysis. He was one of the very first proponents of the idea of an electronic health record in Europe. In fact, Ceusters was based in Belgium, where some of the earliest successful experiments in electronic health record technology were carried out, partly by Ceusters himself. One ontologist told me that Ceusters wrote to congratulate Smith and said “I've been studying your work for some time. I'm really happy that you are now going to start working on medicine. I want to come and work for you.” In the end, Ceusters did go and work with Smith at the institute Smith founded in Germany, and then later in Buffalo. Ceusters’ company was bought by Nuance, an influential speech comprehension company that focuses a lot on medical speech understanding and is part of the IBM Watson Medicine project. Because Nuance bought the ontology that Ceusters built for medical natural language understanding, there is some influence from his work on the IBM Watson project.

Ceusters is Director of the ORG at the CBLS, Director of Research at the Institute for Healthcare Informatics, and Professor in the Department of Biomedical Informatics, all at the University at Buffalo. He is a medical doctor who specializes in neuro-psychiatry. When he was doing what is the equivalent of what is residency in Europe, he earned a master's degree in informatics and later another in knowledge engineering from the Babbage Institute for Knowledge and Information Technologies. He then worked for a brief period of time in the Department of Psychiatry at the University Hospital of Ghent. His specialty was in helping the department with informatics. Owing to his success at Ghent, Ceusters received a few awards and was then invited to join the Department of Biomedical Informatics, which was headed then by Georges De Moor, one of the pioneers of standardization in healthcare informatics. De Moor asked Ceusters to work on a research project that he initiated, acting as principle investigator. Ceusters related to me that he finished that project successfully, although it was quite difficult for him. After the success of the project with De Moor, Ceusters explained to me that he thought that he could write a grant himself, and so he did. He wrote a grant about Natural Language Processing through the Department of Medical Informatics in Ghent and it was accepted. Six months later he wrote a second grant which also got accepted. On the bases of this success, he started his own research company and began to receive massive amounts of support from the European Commission, particularly in the areas of natural language processing (which Ceusters applied to medicine). One of his dreams, he told me, was to “see every health care provider using an electronic health care record system.”

One ontologist explained to me that at the time, the United Kingdom, France, and Belgium were strong countries in the adoption of electronic health care. Belgium was strong for primary physicians but not for specialists in informatics and support. The main reason and problem being that doctors usually do not have the time to fill out the structured information that is actually required to make machines be able to reason with the data. The solution that Ceusters had in mind was to use natural language understanding—doctors could just speak and the machine would transform the language in such a way that it would be able to convert it to structured information. While Ceusters was working on that, he was approached by somebody who said he should bring the technique on the market instead of simply doing research. Ceusters then created the company Language and Computing in 1998, which focused on semantic indexing for medical documents. They developed what at the time became the largest Biomedical Ontology. The company had about forty-five employees; twelve were medical doctors, primarily refugees from South America and Eastern Europe—people who had a medical degree in their home countries but who were not allowed to practice medicine in Belgium. Ceusters gave them a job as content experts, but he also employed computational linguists and software engineers.

It was while doing this type of work that Ceusters eventually met Smith. In 2001, after Smith won the prestigious Wolfgang Paul Award, Ceusters was surprised to learn that Smith had received the prize. Ceusters explained to me (W. Ceusters, phone interview, September 8, 2015) that at that point he had already read a couple of Smith’s ontology papers but he really had no idea who Smith was. "The only thing that I knew was that he was a philosopher," he told me. "I was really pissed off…" (W. Ceusters, phone interview, September 8, 2015). After Ceusters had tried with great difficulty to secure European money to build a medical ontology – he in fact did receive a lot of European money but never for that specific task – Ceusters was upset that Smith – who had virtually no experience in medicine – was awarded such a large prize for that very task. Ceusters wrote to Smith on a Sunday morning via email, and surprisingly Smith immediately responded. They conversed a little bit, picked up the phone. Two weeks later Smith arrived at Ceusters' office in Belgium to see what kind of research he was doing.

> He looked at my system and he said, 'Oh yeah, I mean that's absolutely fantastic but that's wrong and that's wrong and that's wrong.' I said, 'Why is that wrong?' He gave me some reasons that I didn't understand because in those days I was working on what we would call concept-based ontology and I wasn't aware of the ontological theory of realism. Barry introduced me to that, and I am always eager to learn. We made an agreement that I would use his knowledge to understand what he was talking about to make our system better. And I would introduce him to all my connections in healthcare. We formed an extremely good team (W. Ceusters, phone interview, September 8, 2015).

Ceusters related to me that it was through Smith that he became acquainted with ontology and that after that he started to follow the literature, eventually looking back to previous philosophers. "At some point, I came across Kripke, which I didn't think much of honestly because of his possible world semantics, which I think is a nice trick, but is not ontologically well founded" (W. Ceusters, phone interview, September 8, 2015). At some point Ceusters, like the ontologist I spoke to, came across Ingarden, who Ceusters admits did the type of "interesting work" in ontology that he had been following. Around this time, likely in 2004 or 2005, the board of Ceusters' company, which was by that time a seven million dollar company, did not want to invest any more in research because they were already five years ahead of the competition. “They didn't accept my plan for a new research grant,” Cesuters told me. “At that point, I simply left the company and worked with Barry in Germany where I created the Center for Ontological Research at the University of Saarbrücken. When the money from the Wolfgang Paul Award ran out, Smith invited Ceusters to join the newly created CBLS.

Though not connected to Smith and Ceusters’ realist project, Guarino is another figure who is important to the history of ontology research. Guarino is Research Director at the Institute of Cognitive Sciences and Technologies of the Italian National Research Council (ISTC-CNR), where he leads the LOA in Trento. Since 1991, he has played a central role in ontology, emphasizing an interdisciplinary role in ontology-building that combines philosophy with science. Guarino and Poli coedited a 1995 issue of the *International Journal of Human-Computer Studies* on the role of formal ontology in information technology. Guarino helped lead teams to develop the OntoClean methodology and the DOLCE foundational ontology. His Current research interests include service science, socio-technical systems, and e-government. He is founder and editor-in-chief (with Mark Musen) of the journal *Applied Ontology*, founder and past president of the International Association for Ontology and its Applications, and editorial board member of the *International Journal of Semantic Web and Information Systems* and *Journal of Data Semantics*.

3.5 Communities in Ontology History: Conferences, Email Lists, Organizations

Ontology work has largely been coordinated over email lists and through the internet by interdisciplinary teams of researchers. Studying ontologies should involve knowing the history of such groups and how they contributed to the rise of a multiplicity of ontology projects that vied for dominance in the field before the realism-based project took over. The communities in ontology history, their conferences, email lists, and organizations show that computational ontology is a communal endeavor that requires the participation of a multitude of different researchers who have specializations in different fields.

As technologies of individuation that are able to partition reality into computable portions, ontologies necessitate the integration of specialized domains of knowledge that are often far removed from one another. The internet, message boards, emails, and conferences of emerging ontology communities needed to be formed and established with the aim of facilitating research across such disciplinary domains and maintaining longevity. Further, the specialized tools of ontology-based work were developed in communities and had to be shared to enable mentorship and apprenticeship relations where newcomers could learn the difficult new languages and technologies of ontology. The loose knit cultures that formed around ontology work, particularly on the internet and through conferences, allowed for the dissemination of these language and cultures, furthering the ontological enterprise.

An extremely valuable summary of the history of modern computational ontology, including excerpts from transcripts of multiple oral histories of ontology research, is included in Yim (2015). Yim himself is a pioneer in applied ontology and sought to document some of the key communities that formed during the rise of computational ontology as a discipline. Yim describes some of the key events that led to the emergence of international communities structured around ontology discourse and development. In the first presentation of the International Association for Ontology and Its Applications (IAOA), Guarino noted that a 1993 workshop served as the first major event that marked the development of ontology communities (Guarino, Oberle, & Staab, 2009). Yim states that this was one of the first times where many different researchers from multiple countries convened to discuss issues related to applied ontology. The event was the International Workshop on Formal Ontology in Conceptual Analysis and Knowledge Representation and it took place in Padova, Italy.

The Protégé user community is one of the largest ontology communities in the world with over 200,000 registered users. Protégé was developed by the Stanford Center for Biomedical Informatics Research at the Stanford University School of Medicine and is an open source ontology editor that includes a graphic user interface to help define ontologies. Yim notes that Protégé has its roots in work doctoral work of Mark Musen from the 1980s and that Protégé's "community was created around the tool/technology, and served to foster collaborative research, development, education, and user support" (p. 3). It was around 1995 at a conference that Musen recalls gathering with some colleagues at a bar to hold one of their first meetings, eventually settling on establishing an email list. In the 1990s many communities formed around specific domain ontologies dedicated to specific content, such as the Gene Ontology, for example. Yim recalls speaking with the Gene Ontology's creator, Chris Mungall, who noted that various GO mailing lists were established for GO Consortium members and people applying GO terms to genes as part of their bio-curation work. "While the size of the GO discussion list is 100-200 subscribers, the size of the wider community of users is much larger, as there are hundreds of software applications doing GO-based analyses, and each of these has its own community of users (although those communities are no longer tightly knit together)" (Yim, 2015, p. 3).

The Open Biomedical Ontologies (OBO) initiative started in 2001with the aim of coordinating "with other ontology developers for the life sciences so that they would apply the key principles underlying the success of the GO – namely, that ontologies be open, orthogonal, instantiated in a well specified syntax, and designed to share a common space of identifiers" (Yim, 2015, p. 3; Smith et al., 2007). As the most comprehensive ontology repository in the world, the OBO Foundry has one of the largest ontology communities in existence. The OBO Foundry will be discussed in chapter five. Currently, there are hundreds of ontologies in the OBO Foundry and the email list of the Foundry has over three hundred subscribers (Yim, 2015, p. 4). The First Formal Ontology in Information Systems (FOIS) conference was held in 1998 in Trento, Italy. Yim notes that like "other scientific conferences, the FOIS conferences provide the ontology research and development community with a familiar platform for collaboration. FOIS addresses diverse domains, such as conceptual modeling, database design, software engineering, organizational modeling, AI, computational linguistics, the life sciences, bioinformatics, geographic information science, knowledge engineering, information retrieval, and the Semantic Web" (p. 4). The efforts at FOIS are collaborative and involve large groups of interdisciplinary researchers who specialize in applied ontology-making.

An ontology workshop in 1998 (the same year as the Trento conference) held in Heidelberg, Germany brought together diverse groups with an "attempt towards achieving some sort of convergence on basic ontological categories and relations among representatives of a broad interdisciplinary community […] the mail thread among participants of this workshop initiated an informal virtual collaboration of formal ontology researchers" (Yim, 2015, p. 4). Yim notes that the Standard Upper Ontology Working Group (SUO WG) evolved in part in 2000 thanks to the events at Heidelberg and built a community that

> worked and thrived on a set of mailing lists hosted by IEEE […] The SUOWorking Group grew to 88 voting members and logged 25,000 postings. They went as far as shortlisting six candidate upper ontologies, but never came to building a consensus on the one "standard upper ontology" they had set out to create. The SUO mailing list, however, set the stage for virtual collaboration among geographically distributed members of the community, and paved the way for some of the same players to collaborate in the Ontolog Forum, after the Ontolog community emerged in 2002 (p. 4).

In 2002 the ONTOLOG community of practice comes into being led by Jon Bosak (Bosak was widely regarded as the father of XML). Yim suggested to Bosak that he convert his Universal Business Language into an ontology. Subsequently, the UBL-Ontolog mailing list was formed. Even at that time, Yim notes that ontology "clearly wasn't mature enough to deliver what was needed then. Everyone reached consensus that the effort would best be spun off and made into an independent mailing list, which would garner even more support from a broader community. Bosak (who was actually a philosophy major at college) was very supportive" (Yim, 2015, p. 5). The Ontolog community was created by Kurt Conrad, Leo Obrst and Peter Yim 2002 and an additional email list was created to open up ontology to a wider community. Yim describes Ontolog as an "open dialog in ontology" that was "designed to be a Community of Practice (CoP) in the sense that John Seely Brown (of Xerox PARC) would have it: "a small group of people who have worked together over a period of time. Not a team, not a task force, not necessarily an authorized or identified group. They are peers in the execution of "real work." What holds them together is a common sense of purpose and a real need to know what each other knows" (Yim, 2015, p. 5; Brown & Gray, 1995). Yim describes a surge of energy taking place in the ontology community at that time, with many engaging online discussions, activities, and projects. Yim writes that the Ontolog community "grew into, arguably, the strongest community of its kind – an open, international, virtual, community of practice devoted to ontology, ontological engineering, and semantic technologies. Ontolog developed processes that were regarded as community best practices that others were following" (2015, p. 5). Yim retired in 2014 and left the operations of Ontolog to a board of trustees.

In 2006, the Ontology Summits began, organized by Patrick Cassidy, Leo Obrst, Steve Ray, and Yim. The focus was on upper ontology and in the same year the Upper Ontology Summit was held in partnership with the US National Institute for Standards and Technology. The event was an opportunity to bring many different upper ontologies together to converge on key issues. Yim writes that the Ontology Summits were structured around a specific theme each year, allowing individuals interested in ontology to work together and share their research. He also write that the Ontology Summit "process matured as time progressed, and this annual program now comprises almost four months of virtual discourse (over archived mailing lists) and virtual panel presentations and discussion sessions (over augmented conference calls), and culminates in a two-day face-to-face workshop, during which the community, among other things, shares its findings and present its distilled thoughts in a collaboratively developed communiqué" (2015, p. 6). Multiple groups have helped in the organizing of the summits, including the Ontolog Forum (Ontolog), the National Institute of Standards and Technology (NIST), NCOR, the National Center for Biomedical Ontology (NCBO), the IAOA, and the US National Coordination Office for Networking and Information Technology Research and Development (NCO-NITRD). Various themes have included taxonomies and folksonomies, open ontology repositories, ontology standards, ontologist training, Big Data and semantics, ontology evaluation, and the Internet of Things.

The International Association for Ontology and Its Applications (IAOA) started in 2009 and Guarino was chosen to lead the project. The IAOA is an international non-profit organization dedicated to promoting interdisciplinary research and collaboration at the intersection of philosophy and computer science. The main activities of the association as stated on their website include educating stakeholders in the practice of ontology making and in how ontologies can be utilized for a variety of purposes. Promoting interdisciplinary collaboration among private and public organizations, networking with national and international groups, supporting the research and development of ontology in science and industry, supporting personnel exchanges, nurturing ontology learning in developing countries, facilitating the publication of books, journals, and conference, distributing awards and scholarships, and the creation of ad hoc groups to solve new problems. IAOA's main conference is the FOIS and it has been held biannually since 1998. The IAOA "designated the journal *Applied Ontology* (which was started in 2005) an affiliated publication" and "hosts various workshops, operates a range of technical committees and special interest groups, provides scholarships and other incentives to upcoming scholars in the field, and supports other related professional events" (Yim, 2015, p. 7-8). Lastly, there have been many smaller organizations that have been involved in applied ontology standards over the years, working in partnership with organizations such as the International Organization for Standardization, International Electrotechnical Commission, the United Nations Centre for Trade Facilitation and Electronic Business, the Institute of Electrical and Electronics Engineers, the Organization for the Advancement of Structured Information Standards, the Object Management Group or the WorldWideWeb Consortium (Yim, 2015, p. 8). Such standards are discussed in the next half of the dissertation.

Understanding applied ontology through its history and communities sheds light on the emergent nature and progress of ontology from its multiple origins. Qualitative, critical, and historical analysis also provides a look into the network of individuals who might be accountable for the ethical problems that might arise in future applied ontology work. The individuals, groups, and institutions in applied computational ontology illuminate the invisible infrastructures through which ontologies are created and thus enable a closer look into ontological practice and methods. Critical Data Studies must focus on the methods but also the concrete structures through which those methods are developed—it is through a foregrounding of such structures that the black box of ontology can be opened.

CHAPTER 4. BASIC FORMAL ONTOLOGY AND THE FOUNDRY

4.1 Chapter Summary

This chapter uses the theory of technologies of individuation to describe the science teams, editors, rules, methods, and standards (Star & Lampland, 2009) that make up the BFO. The BFO was developed by Smith and researchers at the University at Buffalo and is an upper-level ontology used to organize ontologies that are domain specific. It is the most successful upper-level ontology to date that enables semantic interoperability between domain specific ontologies. The chapter then goes on to describe the Open Biological and Biomedical Foundry (OBO), an online resource that houses domain specific biological and biomedical ontologies that adhere to strict BFO principles. The OBO Foundry is one of the largest collections of domain specific ontologies that are semantically interoperable, thanks to the BFO. Primary data was gathered from interviews with individuals involved in the building and organization of BFO and OBO and secondary data was collected from archival research at the OBO Foundry, including data on the Foundry and its operational policies, including access, management and dissemination, capacity building resources, and protected information. The thrust of this chapter is to show that the BFO and OBO should be viewed as technologies of individuation that enable to organization and sharing of data as representations of invisible partitions in reality.

4.2 The Basic Formal Ontology Discuss Group

In Simondon's philosophy, individuation is meant to be understood as a process of informational individuation whereby entities emerge through levels of informational abstraction and granularity. Technologies of individuation are computational emerging technologies that translate data to sharpen our understanding of the world and the levels of informational abstraction through which it is accessed. Such technologies sift through data disorder to locate complex patterns and meaning across disparate datasets and in doing so they alter our conceptions of materiality. The subject of technologies of individuation is informational ontology, meaning the way in which materiality is rethought along information-theoretic lines. Technologies of individuation re-ontologize reality and produce new realities through informational ontology by using data to understand reality as relational and consisting of abstract levels of granularity. Informational individuation is a process whereby technologies of individuation such as computational ontologies use heterogeneous data to individuate new entities that emerge through the process of semantic interoperability. By achieving semantic interoperability through applied informational ontology computational ontologies individuate entities and locate new forms of reality.

As technologies of individuation involved in the partitioning of reality into representative data artifacts that are semantically interoperable using computation, ontologies first require a team of professionals to create, maintain, and monitor their infrastructures. Similar to the ad hoc communities formed around ontology engineering groups, institutions, and conferences mentioned in the previous chapter, the method of ontological realism similarly requires an exposition of the social lives of the groups and individuals that participate in its everyday activities. Such activities typically begin with establishing a means of communication between the large groups of researchers who must discus BFO principles and interact with each other on a daily basis. Individuals involved in massive group operations in ontology-building must communicate with each other while sometimes being located in different geographic locations. Further, a close look at such groups and their communities of practice illuminate the methods through which decisions are made and discussions are conducted. In the case of the BFO, the public material that has been generated by the community of practice that has formed around applied ontological realism shows that there is a degree to which scientific rhetoric plays a productive role in the communication, formation, and maintenance of BFO methods (Gooch, 1975; Depew & Lyne, 2013). Communities of practice use a variety of tools and techniques to maintain contact, including technologies such as Google Groups and other forms of online social networking. The BFO Discuss group is one such community (Figure 4.1) where members involved in the day-to-day activities of the BFO participate and discusses issues concerning methods, implementation, editing, and infrastructure. There are currently 536 topics in the BFO discussion group which effectively act as message threads that contain many more individual messages. The group was started by Holger Stenzhorn in 2006 to provide a space to enable groups of researchers interested in applied computational ontology a home to work through some of the BFO's theoretical and methodological problems. The group has been the main method of contact for individuals involved in ontological realism using the BFO method.

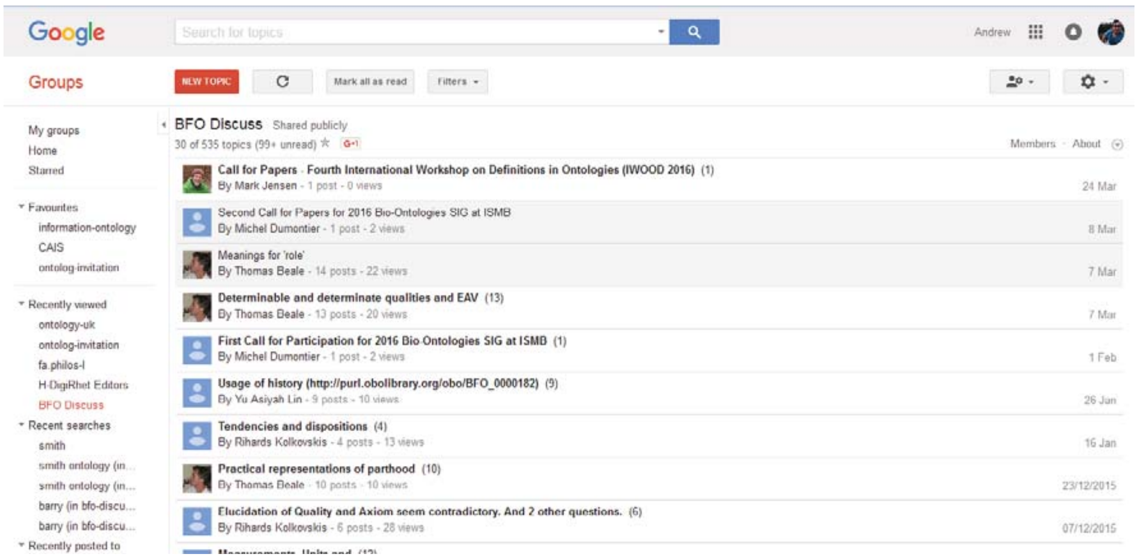


Figure 4.1 BFO Discuss – the Basic Formal Ontology Google Group

Within such groups, members discuss the terminologies that must be adopted by the BFO, including borderline cases in the natural sciences, but also how various types of documents are to be represented in the BFO (Smith, 2014; Almeida, Slaughter, & Brochhausen, 2012). For example, how is a digital document like a prescription to be referenced in the BFO? The BFO Discuss group allows for debates to be aired openly and in the public so that input can be received from any number of individuals. For example, on top of debating scientific entities in fields such as biomedicine, members of the group have debated document types (virtual digital artifacts and virtual transactions) and how such artifacts should be represented in the BFO, including forms, templates, memos, but also things such as entire archives, protocols, amendments, dates, maps, photographs, and diagrams. They also discuss what can be done to documents, including things such as stamping, approving, cancelling, filing, etc. Using the realism-based method of BFO, BFO Discuss members debate how to include things such as funding data and legal actions in the ontology, and also things like errors, forgeries, and invalidity. Institutional systems and entities such as government agencies and nongovernmental organizations must be represented in the ontology. Many of these organizations represented entities that must exist in a social ontology where the method of ontological realism becomes more complicated—social structures do not depend on such things as universals in science, thus arguments in favor of the inclusion of entities in a social ontology are difficult to make. We will return to such difficulties in the next and final chapter, but for now we will turn to the debates that occur in the BFO Discuss group and their significance for the BFO method.

The publicly available BFO Discuss group contains records dating back to October 15th 2006. The first message, written by Stenzhorn (currently Stenzhorn is working at the Universität des Saarlandes) reads,

> Welcome to "BFO Discuss",
>
> I have created this group for discussions pertaining to the theoretical and practical aspects of the "Basic Formal Ontology (BFO)". If you are interested in this particular topic I welcome you to join this group and to actively participate in it.
>
> Regards,
> Holger Stenzhorn

The introductory message from Stenzhorn shows that BFO Discuss was created to be publicly available and open to everyone including specialists and generalists. This follows Yim and Smith’s tenant that ontolgoy work should be open and accessible to the widest variety of users. It is through collaborative effort that ontologies like the BFO are maintained and updated into further iterations. Such collaboration also shows that the scientific realism endorsed by the BFO is still up for debate in such groups where terms can be debated on and everntually implemented or erased. Such a dialogic approach to applied computational ontology rules and methods runs the risk of falling further from claims of scientific objectivity and closer to ad hoc conceptualist approaches. Yet, BFO Discuss contains examples of many practical issues that are unrelated to definitions that still come up in applied computational ontology work, including transfer issues, as the second message in the group indicates. Effectively the first problem posted in the BFO Discuss group, the next message in the group reads:

> Hello BFO gurus.
> I hope this is the right site to discuss some issues that came up during the BFO import from OBI. Taken from the OBI wiki:
> https://www.cbil.upenn.edu/fugowiki/index.php/BFOImportStatus:
> The following OBI top level classes may not belong into OBI and are perhaps better suited to exist in BFO (or PATO)?
>
> 1. OBI:time_interval (duplicate with List 2)
> 2. OBI:time_point (duplicate with List 2)
> 3. OBI:state (duplicate with List 2)
> 4. OBI:characteristic (duplicate with List 2)
> 5. OBI:value
> 6. OBI:cardinal_part_of_value
> 7. OBI:physical_entity (duplicate with List 2)
> 8. OBI:immaterial_entity
> 9. OBI:material_entity
> 10. OBI:boundary (duplicate with List 2)
> 11. OBI:material_entity
> 12. OBI:collection_of_physical_entities (dup with List 2)
>
> Do OBI:physical_entity and material_entity belong under snap:object and immaterial_entity under snap:site ?
>
> Are the following classifications correct?

> obi:state is_a bfo:quality
> obi:characteristic is_a bfo:quality
> obi:time_point maps bfo:temporal_instant
> obi:time_intervall maps bfo:temporal_interval
> obi:boundary maps bfo:object_boundary (and bfo:process_boundary)
> obi:spatial_region maps bfo:site
> obi:collection_of_physical_entities maps bfo:object_aggregate
> obi:population is_a bfo:object_aggregate
> obi:sample_population is_a bfo:object_aggregate
>
> is obi:data_set a bfo:object_aggregate subclass? Is obi:time_interval a bfo:continuant , when in the bfo definition is stated that a continuant can not have temporal parts ? obi:data_set is_a bfo:object_aggregate subclass ?
>
> Any thoughts and comments appreciated.
> Cheers, Daniel Schober

From this beginning stage, the BFO Discuss group was used to ask and answer questions about the methods of BFO and ontological realism in the context of a community of users and specialists who kept up to date with the practice of applied ontology engineering. In the example given above, Daniel Schober asks “Do *OBI:physical_entity* and *material_entity* belong under *snap:object* and *immaterial_entity* under *snap:site* ?” The question concerns how BFO functions once it is populated by the content of another ontology, which in this case is the Ontology for Biomedical Investigations (OBI). This is significant since the OBI is a domain ontology that contains biomedical data that are sensitive and should be carefully approached when being searched and re-categorized according to a new ontological method such as that contained in the BFO. The OBI is a good example of how domain ontologies interact with the BFO to produce new technical problems that can also be potentially ethical problems. For example, in running OBI data through the BFO, the data contained in the OBI have their own anonymization and codes and there is a potential for these to be complicated when being processed through the BFO. Further, complications can arise in the translation of terms from the OBI to the BFO. Schober asks whether or not "*obi:population is_a bfo:object_aggregate*" to see if the translation from the OBI to the BFO principles is correct. The BFO follows much more abstract rules and principles and this is part of the reason it is so successful (notice the transfer of the class *population* in the OBI to the class *object_aggregate* in the BFO). Such a translation of existing data into much more abstract terms of categorization threatens to undermine the existing data structure, especially if the individual who originally defined the data is not present to interact and dsiscuss translation rules before integration. Similarly, in a threaded discussion that runs from February 4th 2013 to April 25th 2013, members of the BFO Discuss group argue over the definitions of processes and entities. Combing through the vast archive of publicly available material on the BFO Discuss list, I came across numerous examples where members of the group engaged in heated dialogue with one another on some of the basic principles of ontological realism. For example, in a post from February 24th 2007, Smith himself had to intervene in a debate that was happening in the discussion board to clarify mistakes concerning philosophical principles in boundary theory and logic (Figure 4.2).

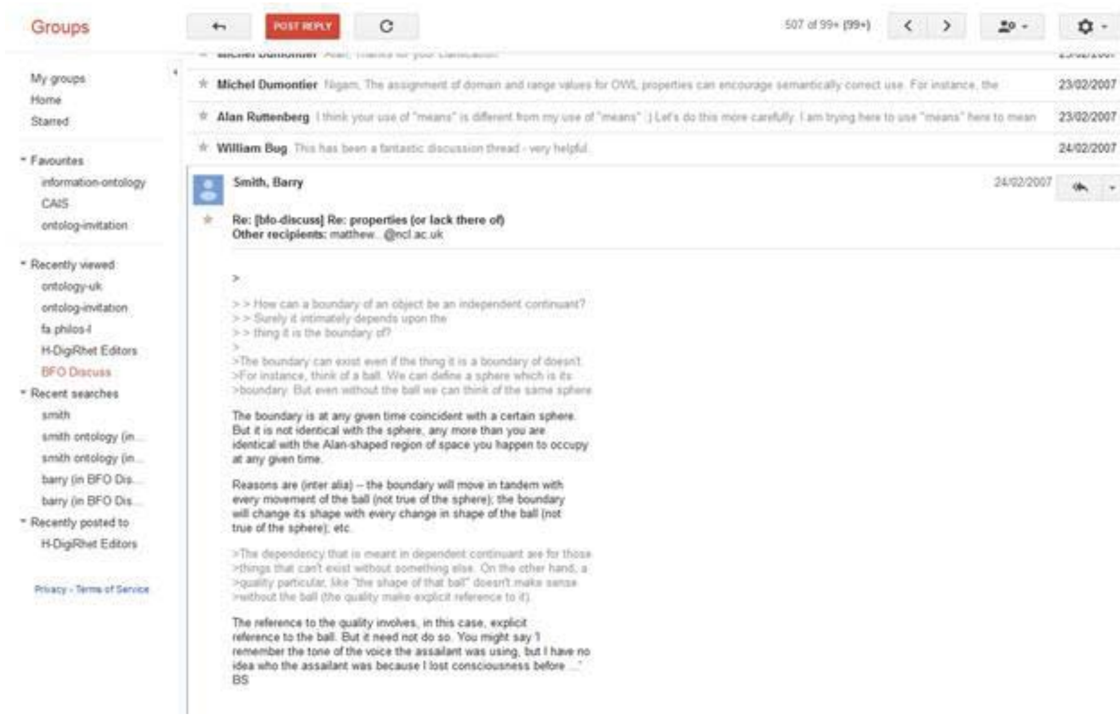


Figure 4.2 Debate in the BFO Discuss Google Group

Members of the BFO Discuss group participate and attempt to answer questions about ontological realism as they come up from a variety of users around the world. Such a group represents a subset of the ontology research community, one that has formed around a specific subject (BFO) using a specific technology (Google Groups) to facilitate community action. As an example of scientific culture forming online, the BFO Discuss group facilitates the "arguments and beliefs to which there is a constant appeal in daily life" of scientists (Latour & Woolgar, 1979, p.55). The BFO Discuss group acts as a venue to extend dialogue around issues that align with different arguments and belief systems that can potentially affect the applied ontological technology since the BFO method involves definitions and relations which are constantly revised and susceptible to biases and beliefs. For example, in Figure 4.2, Smith attempts to reason with the group to explain why *boundaries* should be considered *independent continuants* and how they should be included in the BFO. Alan Ruttenberg, in a previous message, asks “How can a boundary of an object be an independent continuant? Surely it intimately depends upon the thing it is the boundary of? The boundary can exist even if the thing it is a boundary of doesn’t. For instance, think of a ball. We can define a sphere which is its boundary. But even without the ball we can think of the same sphere.” Smith in turn replies: “The boundary is at any given time coincident with a certain sphere. But it is not identical with the sphere, any more than you are identical with the Alan-shaped region of space you happen to occupy at any given time.” In the end Smith convinces Ruttenberg that a boundary is an independent continuant. Such exchanges show that the BFO is subject to debates and exchanges that can potentially affect its method and this suggests that there is some degree of rhetorical force involved in the construction and communication of the BFO itself in terms of its scientific principles (Gooch, 1975).

BFO Discuss is not official but is rather an informal ad hoc community formed around a recognized need to provide support to the ontological realist project. Like the many other historical ontology discussion groups noted by Yim (2015), BFO Discuss is the latest iteration in a line of communities that has formed around the practice of applied ontological engineering. Such virtual communities (Rheingold, 1993) contribute to transparency and add to the overall value of the realist ontological project by enabling individual’s direct access to knowledge bases and specialists whom they might not have access to otherwise. Historically, such groups are familiar in the early development of emerging technologies—the internet and email being just two examples. Similarly, as Yim (2015) has shown, ontology work since the 1990s has embraced a somewhat inclusive and community-oriented nature owing to the large amounts of interdisciplinary resources that are required of applied ontology work. There is a certain sense in which applied ontology work does not fall under any specific domain and that it is the result of numerous forces of activity and specialized knowledge that form in ad hoc communities. Computational ontologies are in this sense like the alleged black arts that consist of techniques or practices that are mysterious or sinister—the practice of applied ontology work is somewhat opaque and hard to pin down, yet applied computational ontologies as technologies of individuation as extremely powerful. A black art that engages interdisciplinary and multidisciplinary fields, computational ontology elicits work from philosophers, logicians, engineers, computer and information scientists, managers, community-builder, and many more.

The openness and community-centered work of the BFO ensures access and increased accountability in the ecology of information sharing in the digital sharing of ontology resources, commentary, and critique (Christen, 2009). BFO Discuss represents the type of shared community actions that coalesce around problems in the application of ontological realism. What is interesting to note is the sheer number of individuals who are involved in the discussions and decision-making process. Smith (who is for better or worse the lead inventor of BFO) himself has engaged in discussion regularly, as have numbers of rank and file engineers, computer scientists, even other philosophers. The discussions are often illuminating and provide a snapshot of the rhetoric that ontologists engage in when arguing about ontological realism and its principles, suggesting that there is a degree to which scientific rhetoric (Depew & Lyne, 2013) plays a role even in the methods of scientific ontological realism.

4.3 I'd Rather Share My Toothbrush (Referent Tracking)

In *Life out of Sequence: A Data-Driven History of Bioinformatics*, Stevens (2013) describes that when he "raised the subject of ontologies among biologists" they often responded to him with a quote that is sometimes attributed to the influential biologist Michael Ashburner. Ashbrunner is reported to have said that "biologists would rather share their toothbrush than share a gene name" (Stevens, 2013, p. 125). This type of guarded cloistering is a regular enemy in applied ontology work—researchers would often rather share their toothbrush than share names in an ontology. The reason for this is that researchers develop and use their own naming procedures that have usually been in place for a substantial length of time. Such procedures are often clumsy and not well-defined yet the length of time that research teams have spent with the data make it difficult for them to envision seeing such data be integrated with others if such an integration were to require a substantial change in their referent tracking (the names for data artifacts). It is a bit like already having bought a vacation to less desirable local when a newer, better, yet more expensive option is presented. People will likely stick to the original destination (it was good enough to begin with) rather than go through the process of cancelling their flights, hotels, and other travel plans to rebook everything for the new destination. In fact, much of the time the problem of referent tracking occurs even before the issue of semantic interoperability via applied ontology is relevant. Data are simply not categorized in an accurate fashion, or individuals have strong preferences for how they would like to label their data, resulting in heterogeneous data structures that do not cohere.

Part of good computational ontology work includes good referent tracking before data are even run through the ontology (recall Ceuster's example from the first chapter). Computational ontologies and ontologists produce what Coleman (2010) has described as "vernacular cultures" in digital media, meaning the special types of languages, codes, naming systems, and ways of labelling that develop when specialized fields begin to amass massive amounts of data (the BFO similarly adopts a vernacular even though it is given upper-level ontology status—it is not immune). Such vernaculars extend to domain-specific ontologies, data sets, and researchers as well and are represented in the various types of referent tracking that data scientists engage in when acquiring data that act as representational artifacts. Data must be tagged, labelled, and stored for the data to be about and represent an entity and such processes often lead to problems in terms of the standards and clarity of data organization. The integration and semantic interoperability of heterogeneous information resources arising from different branches of science is one of the most pressing problems of applied ontology.

Information comes in so many formats – including everything from data about instruments, standards, products, protocols, instructions, codes, and even whole domain-specific ontologies themselves – which are incompatible and formalized only locally. Descriptions are often made using very general terms, are based on natural language, or are adopted according to a formalized classification system. Yet very few naming procedures contain explicit and accurate references to actually existing entities that correspond in a realist tradition. Such a lack of explicit reference is not typically a big issue for normal everyday life in the laboratory; a researchers discuss data during meetings and in everyday parlance are able to share with each other the definitions that are applied to data, and if they are not informed a simple email will correct the lack of knowledge. Humans can easily explain to each other what data are about and interpret their meaning—they are able to disambiguate the reference in general statements and terms by embracing contexts, times, places, and people. Such understanding for machines is much harder and makes achieving the semantic interoperability of data much harder. Typically, even well-formalized data – data that use controlled vocabularies – are still poorly organized due to the adoption of codes or terms that are formal but that do not provide an accurate alternative to general terms in natural language. The errors that can evolve are similar to those that come from description logics (recall unicorns); the outcome might be highly formal but realistically inaccurate an ambiguity issue that has long existed in the field of natural language processing.

Ceusters explains the problem in referent tracking in the following way: Is it possible for different ontologists to produce a collection of data independently from each other that refers to the same portion of reality in a semantically interoperable way? Ceusters has built a successful and influential career in referent tracking working with the ORG at the CBLS in the University at Buffalo and has developed a comprehensive theory around it, specifically in the domain of health and the medical sciences—an earlier influential text in this regard is Rector, Nolan, and Kay (1991). The idea of referent tracking was first introduced in Ceusters and Smith (2005) which explained that the practice would be a new paradigm for data entry and retrieval in electronic health records. The text introduces some of the problems that occur when comparing general terms in a health database to specific terms on the side of the patient and explains how ambiguities can be avoided by referring to such specific terms by way of identifiers instead of codes. Ceusters and Smith's approach to referent tacking follows a realism-based method in that it places emphasis on avoiding concepts as descriptors wherever possible and instead using a realistic approach to name entities using a distinction between three levels of reality: reality, our understanding of reality, and our representations of reality—a thesis that bears a striking resemblance to Floridi's conception of information *as, for,* and *about* reality (2011). The difference between realism-based reference tracking and ontology work is that ontology is concerned with general terms in reality while reference tracking refers to the specific entities that populate the ontology. Thinking back to Husserl's separation between material and formal ontology, referent tracking would fall on the side of the material in that it deals with specific items that can be identified.

One of the big challenges in referent tracking (as in ontology work) is the ability to represent entities that are missing, such as fissures, cracks, crevasses, holes, and tears. A conceptual approach to referent tracking would simply posit that such a claim as "does not have diabetes" is a concept and leave it at that, however this does not present a great deal of accuracy. To account for this problem, Ceusters and Smith introduce the notion of *lacks* into their referent ontology as a relational category that holds between particulars and universals (Ceusters, Elkin, & Smith, 2006)—a rather late addition considering the field of applied computational ontology was already at this point almost two decades old. Another challenge has to do with keeping track of entities over time and their changes in reality, understanding, reassessment, and mistakes. Ceusters and Smith (2006a) introduce the notion of versioning in an ontology to address these concerns and emphasize the notions of class, time, and history.

The relevant applications of referent tracking outside of medical contexts are numerous, including decision support in the semantic web (Ceusters & Smith, 2006b) and digital rights management (Ceusters & Smith, 2007). Interestingly, an article by Ceusters and Manzoor (2010) entitled “How to Track Absolutely Everything” describes using systems of referent tracking implemented in networks for intelligence agencies to help achieve the aims set out in Office of the Director of National Intelligence John Michael McConnell’s *Vision 2015: A Globally Networked and Integrated Intelligence Enterprise* (2008). Ceusters and Manzoor write that referent tracking “uses a system of singular and globally unique identifiers to track not only entities and events in first-order reality, but also the data and information elements that are created to describe such entities and events in information systems. By doing so, it meets the requirements of the Nation’s *Information Sharing Strategy*” (p. 13). In Manzoor, Ceusters, and Smith (2009), a method is presented that enables storage of the contents of Joint Battle Management Language messages in a referent tracking system (Figure 4.3).

```
<OrderPush>
  <Task> <GroundTask>
    <TaskeeWho>
      <UnitID>2TF A TEAM</UnitID>
    </TaskeeWho>
    <What><WhatCode>MOVE</WhatCode></What>
    <Where>
    <AtWhere>
      <JBMLAtWhere>
        <WhereValue>
          <WhereLocation Sequence="0">
              <GDC>
                  <Longitude>48.9100583</Longitude>
                  <Latitude>39.9570562</Latitude>
                  <ElevationAGL>0.0</ElevationAGL>
                </GDC>
              </WhereLocation>
              <WhereLocation Sequence="1">
                <GDC>
                  <Longitude>48.9056792</Longitude>
                  <Latitude>39.9891024</Latitude>
                  <ElevationAGL>0.0</ElevationAGL>
                </GDC>
              </WhereLocation>
            </WhereValue>
          </JBMLAtWhere>
        </AtWhere>
      </Where>
    </GroundTask></Task>
</OrderPush>
```

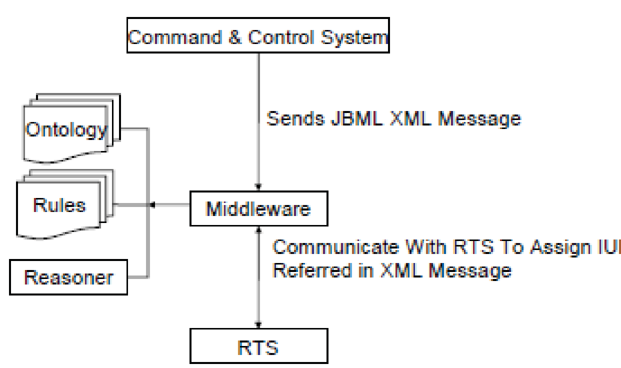


Figure 4.3 Military Referent Tracking (Manzoor, Ceusters, & Smith, 2009)

These and other examples of referent tracking and ontology work in the intelligence world will be covered in the next and final chapter on CDS and ethics. It is in these areas that the realism-based method might encounter some ethical problems, since they concern entities that exist outside of the domain of science, such as human subjects and social groupings.

We have seen how debates occur online in communities such as BFO Discuss and how these discussions can potentially influence the methods and theory of the BFO, suggesting that rhetoric plays some role in the communication of scientific methods and principles endorsed by the BFO (Gooch, 1975). Smith and creators of the BFO must appeal to BFO Discuss members to have them agree and accept their terms and definitions. The introduction of social ontology to BFO represents yet another significant step away from the alleged purely scientific nature of the BFO and points to potential ethical problems in applied computational ontology work. As technologies of individuation that facilitate the processing of informational ontology, applying social ontological entities to the BFO expands the types of entities to which BFO principles are allowed to refer, moving beyond mere objective scientific categorization and into the realm of government control and social engineering (Raskin, Taylor, & Hempelmann, 2010; Mouton, Leenen, Malan, & Venter, 2014). Social ontology can be understood as a type of social engineering in that it is related to information security and can influence the psychological states of individuals who fall within the domain of the ontology (economic actors in an economic ontology, civilians and combatants in a military ontology, customers in a bank ontology, etc.). A concrete example of this is presented in the form of the Military Ontology in the next chapter.

In the effort to track absolutely everything for use in computational ontology, such efforts must be open to critical and ethical reflection. Just as the internet developed as a government sponsored technology fueled by military interests, applied ontology work now – though originating in the world of AI and maturing in natural sciences such as medical informatics – is being applied to spheres that exist outside of science such as the military, but also finance and management. As technologies of individuation that divide reality into invisible partitions according to levels of informational abstraction, computational ontologies that embrace ontological realism must be open to investigation when they are applied to the social world.

4.4 The Method of Individuation: SNAP and SPAN

Opening the black box of ontological realism and technologies of individuation is required to see how computational ontologies might affect individuals, groups, and society. Referent tracking has shown to be an important part of the ontology-building process but how exactly are entities defined at the beginning of that approach and in ontologies? What are the specific methods and conceptual frameworks that are used to define entities for use in referent tracking and ontology-making? Chapters one to three introduced computational ontological realism and tracked its growth out of philosophy into a full-blown network of scientists and researchers who embrace the realist method to advance scientific research. Formal ontology was defined as separate from formal logic and materiality, and first-order logic with universals was shown to be the method of ontological realism. The significance of levels of informational abstraction was discussed along with ontological commitment. Opening the black box of computational ontological realism requires going further and analyzing how the labeling of entities actually occurs during the practice of applying the realist method. What are the exact procedures that are employed when making decisions about how to classify entities for computational ontological realism?

As technologies of individuation, computational ontologies concern the granularity of ontology and the ontological zooming in and out of levels of informational abstraction which are complicated buy the notions of space and time. Spatiotemporal reasoning has long been one of the most difficult aspects of both philosophical and applied ontological work—even fairly recently the topic was not seriously considered in basic applied ontology tutorials (Noy & McGuinness, 2001). The main theory of the BFO has been developed in a series of papers by Smith and Pierre Grenon. One of the most significant theoretical contributions to computational ontology made by Smith and Grenon was the ability to capture time and space in ontological reasoning through a distinction between what they (2004) have called SNAP and SPAN entities. SNAP entities represent a spatial view of ontology supporting snapshot views of reality at successive instants, while SPAN entities represent a spatiotemporal ontology of change and process (p. 137), and it is here where the influence of Ingarden's work on applied ontological realism is most apparent. SNAP and SPAN represent the groundwork from which the BFO was constructed—the realism-based computational ontology that is the BFO grew out of the ideas presented in the SNAP and SPAN theories. Grenon and Smith (2004) develop a theory of modular ontology of the dynamic features of reality and argue that a dynamic spatial ontology "must combine these two distinct types of inventory of the entities and relationships in reality" (p. 137).

Drawing on the work of Ingarden and Zemach (1970), Grenon and Smith (2004) state that the BFO acknowledges that there are *continuants* (entities that preserve their identity over time through changes) and *occurrents* (processes, events, activities, change). Entities persist, while occurrents are bound in time. Zemach’s (1970) concise and power paper entitled “Four Ontologies” seems to have been a rather large influence here. In that paper, Zemach offers the following logical axioms for good ontological work according to a spatiotemporal matrix:

A.

1. At one time a thing cannot be as a whole in different places.
2. At different times, a thing can be as a whole in one place.
3. At any time, a thing must have all its parts in different places.
4. At all times, a thing need not have all its parts in one place.

B.

1. In one place, a process cannot be as a whole in different times.
2. In different places, a process can be as a whole in one time.
3. In any place, a process must have all its parts at different times.
4. In all places, a process need not have all its parts at one time.

Appling such axioms in their description of the SNAP and SPAN method in BFO, Grenon and Smith seem to offer that reality is essentially dynamic but that it is expressed through time. Here, time should be understood in the sense that Ingarden describes and that Smith and Grenon have inherited, meaning that time should be thought of in terms of the duration or cessation of entities, events, and processes at multiple levels of granularity—time emerges out of these structures. Traditionally, philosophical theories tend to privilege either the belief that time and space exist separately and that they have rigorously defined axioms (the snapshot view of the universe) or rather that there is a single overarching spatiotemporal theory (the process view). Grenon and Smith write that what they offer is “a theory that is designed to do justice to what is of value in both of these approaches” and that their “position is that a good ontology must be capable of accounting for spatial reality both synchronically (as it exists at a time) and diachronically (as it unfolds through time), but that these are two different tasks” (p. 137-138). Figure 4.4 is a visual representation of the hierarchy of entities as they exist in the SNAP and SPAN ontology.

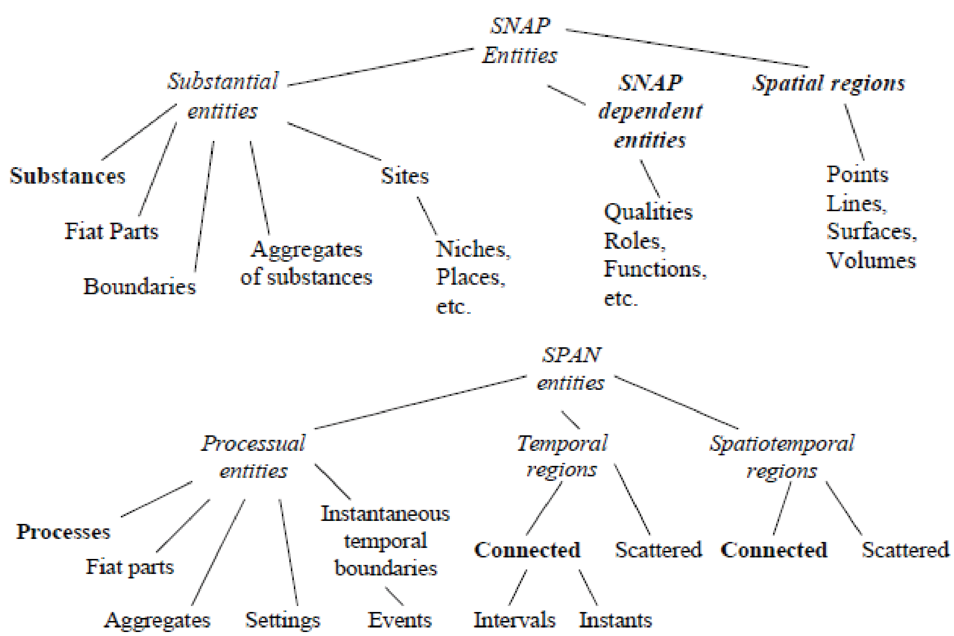


Figure 4.4 SPAN and SPAN Entities (Grenon and Smith, 2004)

Grenon and Smith describe the SNAP and SPAN method as *realist*, *perspectivalist*, *fallibilist*, and *adequatist*. The theory is realist in that they claim it exists separate of linguist or cultural representations, perspectivalist in that they maintain there can be alternative yet equally legitimate perspectives on reality, faillibilist in that it accepts that both theories and classifications can be subject to revision, and adequatist in that it is opposed to reductionism and the view that there is only one basic view of reality (p. 138).

Added to this theory are *fiat objects*, which Smith (2001) describes as boundary objects owing to acts of human decision. In that same way that there are entities that exits – which Smith refers to as *bona fide* objects – fiat objects exist equally according to logically veridical measures. On Smith's theory, all true empirical judgments can be viewed as effecting a division of reality in fiat fashion

> in such a way as to mark out a certain truthmaking region consisting of those entities that are relevant to the truth of the judgment in question. Truth itself can then be defined as the relation of correspondence between a judgment and its corresponding truthmaking region, in such a way that a true judgment would be something like a map of the corresponding portion of reality (2001, p. 17).

Smith's claim that fiat objects represent boundary objects that can be veridical with regard to the way in which those fiat objects are empirically observed in reality echoes Quine's claim about ontological commitments and the bound variables that must exist for the theory to be true. Take the cuts of meat or the map of the United States represented in Figure 4.5. Both of these representations refer to specific entities in reality but are partitioned according to fiat. That the entities (the cow and the landmass of the USA) are partitioned by fiat should not hide the way in which such fiats are *tied* to empirical observations of the reality of those entities. The *chuck* as fiat object corresponds to an empirical boundary in the cow in the same way that *Indiana* as fiat object is tied to an empirical boundary in the USA. Remembering Quine, all that is necessary is that a theory "is committed to those and only those entities to which the bound variables of the theory must be capable of referring in order that the affirmations made in the theory be true" (1948, p. 33). Indeed Smith references Quine in the SNAP and SPAN project and writes that Quine is responsible for coming up with these sorts of fiat objects. The fiat objects *chuck* and *Indiana*, as they exist in an applied ontological framework, need only to refer back to the relation of correspondence that is established in the empirical observation of a portion of reality.

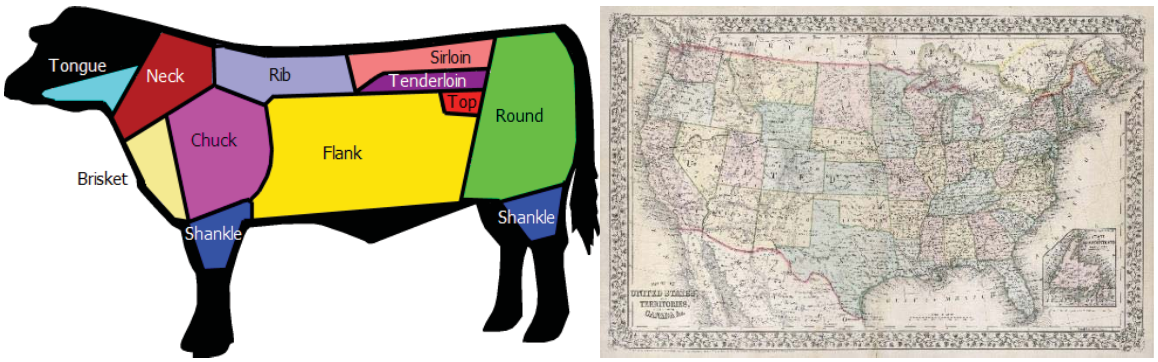


Figure 4.5 Cuts of Meat and a Map of the United States

SNAP and SPAN are the method by which ontological realism functions in applied computational ontologies. They represent the BFO method that informs computational ontology as a technology of individuation in that they are responsible for partitioning reality into levels of granularity and informational abstraction. In this way, reality is presented as a structured informational ontology that can be parsed using the methods of ontology and individuated further and further into multiple entities, attributes, relations, and processes. Figure 4.6 lays out the distribution of SNAP and SPAN entities with fiat objects.

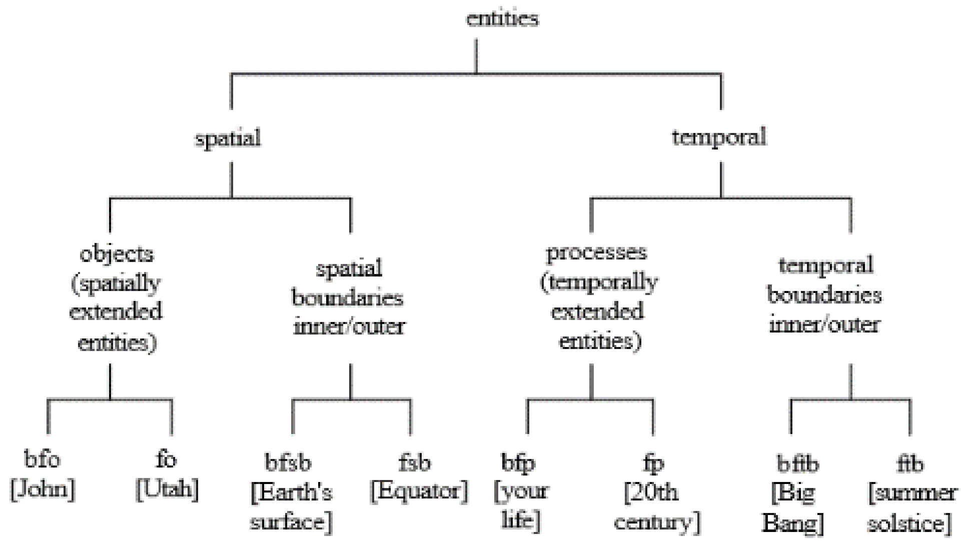


Figure 5. Taxonomy of Fiat and Bona Fide Entities (bf = bona fide, f = fiat, o = object, p = process, sb = spatial boundary, tb = temporal boundary)

Figure 4.6 Taxonomy of Fiat and Bona Fide Entities (Smith, 2001)

Entities are divided into spatial and temporal boundaries. From there, spatial entities are divided into objects that are spatially extended entities and spatial boundaries such as inner/outer. Temporal entities are divided into processes that are temporally extended entities and temporal boundaries such as inner/outer. After these initial partitions, entities can begin to be populated within the ontology. A *bona fide object* is something like John. A *fiat object* is something like Utah. A *bona fide spatial boundary* is something like the Earth's surface. A *fiat spatial boundary* is something like the equator. A *bona fide process* is something like your life. A *fiat process* is something like the 20th century. A *bona fide temporal boundary* is something like the Big Bang. A *fiat temporal boundary* is something like the summer solstice. The methods of ontological realism as expressed in computational ontology partition reality according to such levels of informational abstraction.

Computational ontologies are technologies of individuation that lay bare portions of reality and are responsible for individuating entities so that vast sciences, methodologies, and technologies can cohere using standardized representative data artifacts. The SNAP and SPAN method of BFO is the normative tool that produces semantic interoperability between disparate datasets and it produces new informational ontologies. SNAP and SPAN create an upper-level ontology that semantically constrains the definitions that can be given to data in terms of the reality of time and space. By integrating heterogeneous data using the BFO method, new entities become individuated that emerge from disparate data, producing a new type of informational ontology. The BFO is a technology of individuation that is responsible for making new orders of reality apparent by applying constraining-affordances such as SNAP and SPAN to datasets.

4.5 Basic Formal Ontology, Upper Ontologies, and Domain Ontologies

As I already mentioned, Arp and Spear are coauthors along with Smith of the book *Building Ontologies with Basic Formal Ontology* and they are philosophers. I asked Smith how this team came to be created and he told me that when IFOMIS was founded he advertised amongst philosophers. Smith was disappointed by the degree to which his philosophy colleagues had taken the opportunities created by applied ontology seriously. He still, at that stage, had some optimism and so advertised positions for philosophers in the institute which he had founded. One of the successful applicants was Andrew Spear. Spear was there when they were working on BFO and he created a manual for BFO which was an early version of the book. He was also very much responsible in creating the book itself which is a heavily revised version of his manual plus chapters dealing with things like users and also the formal developments which had taken place since the manual was written at some point around 2005. When the NCBO was founded, Smith advertised again for a postdoc who would work with him on the NCBO dissemination activity. Their job was to disseminate the work of the NCBO including disseminate good ontology practices to people who did not know about ontology. During this period, Robert Arp and Smith worked on various publications which can be seen now as being also extensions of Spear's manual. Some of that work then was incorporated into the book and Arp played an important part along with Spear in putting the final versions together.

Before the development of the BFO and when Smith was working in Italy, the people he was collaborating with had created a somewhat similar upper ontology called DOLCE, which stands for Domain Ontology for Language and Cognitive Engineering. DOLCE was the first upper ontology to be developed and was innovative. The head of DOLCE, Guarino, is somebody who played a historical role in creating the new science of ontology (Guarino and his team are discussed in the first half of this dissertation). Smith immediately saw that DOLCE was important and worked with the DOLCE team. However, there were certain problems with DOLCE. Simplifying, DOLCE was created by people whose native language was Italian and while the ontology was very good, the documentation was less than good. Smith worked with them to try and improve their documentation. For some reason, this collaboration did not work. It was not possible for Smith, as an outsider, to help what he saw as being the primary shortcoming of DOLCE to be rectified by helping them have better documentation. Smith thought that was a sad thing but that forced him to think about how he could create a better upper ontology. When Smith moved to Germany he worked with colleagues there to create what became BFO. Smith's most important collaborator in that connection was Grenon, a former employee of a company called Cycorp. Cycorp was created by an AI guru by the name of Doug Lenat who grew out of the movement to create a formal representation of common sense knowledge. Cyc is an ontology-like artifact which is designed to capture, in first ontologic terms, the entirety of human common sense. Cyc has been funded over the years in part by military and CIA type funding because people think that a resource like that could be useful, for instance, to support natural language processing.

In the preface to *Building Ontologies with Basic Formal Ontology* (2015), Arp, Smith, and Spear write that even though they were professionally trained as philosophers, "What follows is not, however, intended as a contribution to philosophy. It is intended, rather, to form part of what we conceive as the rich, new technical discipline of ontology" (p. x). The BFO is a small, upper-level ontology that is used for organizing domain ontologies and provides the abstract rules and grammar that organize the particulars that populate specific domains. It is the most widely used upper-level ontology in existence today and is used to coordinate the Open Biological and Biomedical Ontologies Foundry (OBO Foundry), one of the largest collection of coordinated ontologies in existence. As such, the BFO does not admit any specific references to material that would belong in domain ontology—thinking back to Husserl's separation of formal ontology from materially, BFO exists on the side of formal ontology and contains only those entities that are abstract enough to satisfy the organization of entities and their relations. BFO emerged from a philosophical project which is focused on the task of providing a genuine upper ontology for domain ontologies that are used in scientific research (such as those in the OBO Foundry). The BFO does not contain references to specific entities in any of the fields that utilize it—there are no references to genes, particles, files, or chemicals, but rather only the definitions of their potential hierarchical relationship.

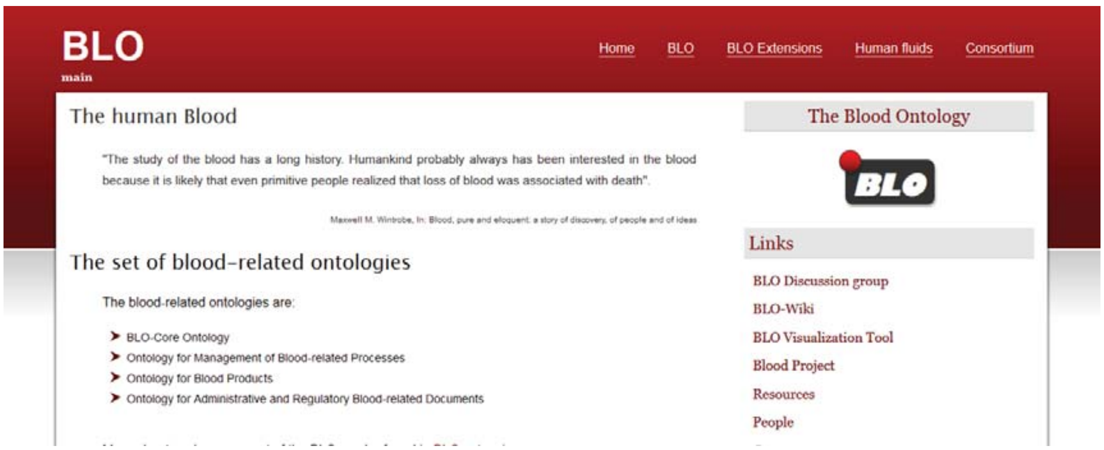


Figure 4.7 Blood Ontology Website

A domain ontology is a representation of the types of entities, attributes, and relations that exist in a specific sphere of reality such as geography, economics, or law. In this sense upper-ontologies are closer to philosophical ontologies since they can be applied to all domains (recall Husserl's distinction between material and formal ontology). Domain ontologies use controlled vocabularies to label data to make it searchable and useable. The Blood Ontology (Figure 4.7) is an example of a domain ontology, since it provides a taxonomy for describing blood-related artifacts and makes communication. Domain ontologies benefit from basic research in formal ontology since it is the task of formal ontology to help in facilitating communication between and among domain ontologies by providing a common language and framework for ontological reasoning.

The BFO (Figure 4.8) was started in 2002 though the original theory was developed by Smith and Grenon in a series of papers, including the aforementioned SNAP and SPAN papers. Since that time various other individuals have contributed in a big way to the BFO project, including Ceusters, Chris Mungall, Fabian Neuhaus, Melanie Courtot, Holger Stenzhorn, Alan Ruttenberg, Mathias Brochausen, Bjoern Peters, Randall Dipert, Janna Hastings, Darren Natale, James Overton, Ron Rudnicki, Stefan Schulz, Selja Seppälä, Jie Zheng, Kerry Trentelman, and over one hundred other individuals involved in the BFO Discuss group. Much of the work has been conducted in Buffalo. There is a team of people lead by Alex Diehl (Assistant Professor in the Department of Neurology at the University at buffalo) which is doing work on neuro ontology, neurological disease ontology. As previously stated, BFO's development has been supported by the Forms of Life project which is sponsored by the Volkswagen Foundation and the Wolfgang Paul Program of the Alexander von Humboldt Foundation. BFO is used by over 130 ontology-driven projects throughout the world. Some of these include the Alzheimer Disease Ontology, Adverse Event Ontology, Actionable Intelligence Retrieval System, Bacterial Clinical Infectious Diseases Ontology, Bank Ontology, Beta Cell Genomics Application Ontology, BioAssay Ontology, Bioinformatics Web Service Ontology, Biological Collections Ontology, Biomedical Ethics Ontology, Biomedical Grid Terminology, BioTop, BIRNLex, Blood Ontology, Body Fluids Ontology, Bone Dysplasia Ontology, Cancer Cell Ontology, Cancer Chemoprevention Ontology, Cardiovascular Disease Ontology, and many others.

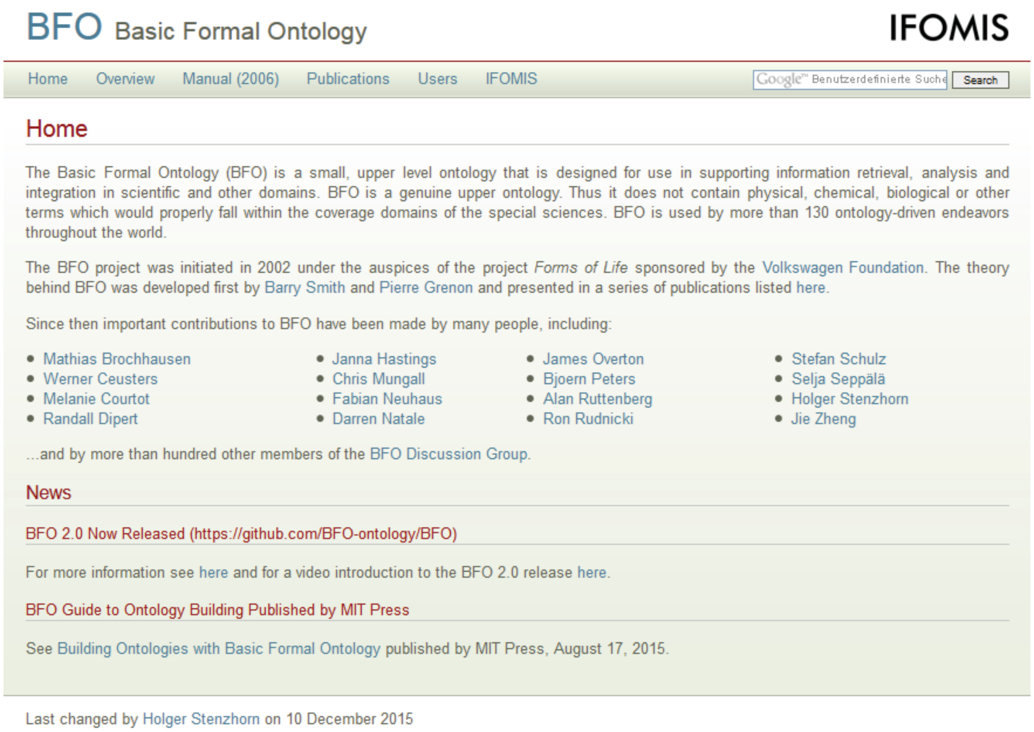


Figure 4.8 Basic Formal Ontology Website

A descendant of Husserl's formal ontological distinction, the BFO is made up of a series of perspectives on reality represented by SNAP (a series of snapshot ontologies indexed by time labeled *Oti*) and SPAN (a single ontology labeled *Ov*). Otis represent an inventory of all entities that exist at a given time. Ov is an inventory of multiple processes that unfold over time. Such processes are invisible in the Oti view while substances are invisible in the Ov view. Depending on the level of granularity and level of informational abstraction, both a SNAP and SPAN method serves as the basis for a series of sub-ontologies. The main outcome of this approach – producing multiple subontologies using the overall method of SNAP and SPAN in the BFO – is that a single portion of reality can be partitioned according to various levels of granularity; entities at one level may be aggregates at a different level. BFO is a technology of individuation that captures the essence of informational ontology and is able to lay bare the invisible structure of the world. As a technology of individuation, what might be seen as a single process at one level might be a part of a longer process at another level a la Ingarden; every ontology is a representation of a portion of reality into Aristotelian categories or universals. In the BFO, ontologies are like windows of a portion of reality; they are partial representations of reality from a specific viewpoint (the ontology). The abstract categories in SNAP and SPAN are meant to be populated by concrete entities in the ontology, entities such as *in the stomach*, *the north of France*, or *below the table* are tied to specific empirical boundaries (they are ontologically committed). Such commitments are bound portions of space.

One important domain ontology related to BFO is the Information Artifact Ontology (IAO). IAO was developed by Smith and Ceusters and is a domain ontology for accurately representing information artifacts. I asked one ontologist about the IAO and they said that they invented it because they "started to think very seriously about what data and information actually is and how it relates to reality. We needed to give a good ontological description of what elements contribute to data and information and so on…That's the part of the information artifact ontology." Ruttenberg is somebody in Buffalo with whom Smith has worked a great deal on BFO, OWL, and IAO. IAO is an attempt to address the following general problem. The domain of the BFO is everything, or at least everything which is a part of the world of empirical measurement. As such, there are no numbers in the BFO ontology. Maybe there should be and maybe there will be—BFO is, like every ontology, always a work in progress. Now, OB is the ontology for biomedical investigations. There is no ontology for investigations generally, since funding has been primarily for biomedical issues and so Ceusters was involved in creating the ontology for biomedical investigations. “Now, we have a problem,” the ontologist said. Biomedical investigations yield information artifacts like databases, publications, footnotes, protocols, and so forth, but these information artifacts are not restricted to the world of biomedical investigations. Exactly the same information artifacts exist in a physics experiment or a report about some kind of human intelligence observation of people moving around in the backstreets of Baghdad. The ontologist decided to create the IAO with the idea that people were using the word “information” in a way in which was primarily influenced by Claude Shannon and Warren Weaver and seeing information as a “kind of jelly.” The ontologist stated that they wanted to focus on the information artifacts which were always created by human beings and that they did not want to get bogged down on issues such as “Is the genome an information object?” They wanted to think about boring things like publications, databases, receipts, bills, orders, licenses, and so forth, because they needed those terms to describe the process of carrying out an experiment.

“You need a license to use a satellite,” one ontologist told me. “Or whatever it might be. We found out that there was no way of fitting in that kind of entity into BFO as it then existed.” BFO has been updated in new versions since its inception. At the time of inventing the IAO, BFO was in its first iteration. The ontologist said that they were “forced to add a new branch to BFO in order to deal with information artifacts, and so we created BFO 1.1. We are now up to BFO 2.0. We've had, so far, three versions. BF0 2.0 is really just a tidied up version of BFO 1.1. The ambitious goals that we had for BFO 2.0 have now been postponed for BFO 2.1. The IAO is not a component of the BFO 2.0. It is its own separate ontology." The problem is that for every entity in any ontology there is a BFO top level node and there was no top level node which would suit license protocol database data entry and so the ontologist's team had to create one that was domain neutral. It just so happens that gene sequences are also subtypes of this BFO category, so it is neutralized between different disciplines—they go into BFO but then the information entity goes into the IAO. Gene sequences go into the sequenced ontology, which is a third, separate ontology from the BFO and the IAO.

Figure 4.9 shows the IAO terms and their definitions. IAO has become a successful domain ontology that is beginning to be integrated into upper-level ontology infrastructure and likely will be included in the next version of BFO. Its importance lies in the way that IAO clearly defines informational objects and digital entities so that they can be included in ontologies without confusion, ambiguity, or error. Before the IAO, there was no stand way to refer to digital objects in an ontology. The IAO provides the semantic interoperability needed to integrate heterogeneous information entities.

Term	Definition
Information Content Entity (ICE)	an *entity* that is *generically dependent* on some artifact and stands in relation of *aboutness* to some *portion of reality* [4]
Representational Artifact (RA)	an ICE which is believed to *represent* a *portion of reality* external to the representation (modified from [5])
Representational Unit (RU)	a **RA** which according to the structural conventions it is designed, is not built out of any other **RAs**
Denotator	a **RU** which *denotes* directly an *entity* without providing a description [6]
Term	a **RU** which is a general expression in some natural language used to refer to *portions of reality* (modified from [5])
Composite Representation	a **RA** built out of constituent sub-representations as its parts (modified from [5])
Data Collection	a **composite representation** built out of **measurement data**
Data Dictionary	a **composite representation** describing, inter alia, what **data items** in a **data collection** are *about*, including a **data format specification**
Terminology	a **RA** consisting of **terms** (modified from [5])
Ontology	a **RA** comprising a taxonomy as proper part, whose **RUs** are intended to designate some combination of *universals*, *defined classes*, and certain *relations* between them [3]
Realism-based Ontology	an **ontology** built out of **RUs** which are intended to be exclusively about *universals* and certain *relations* between them, intended to mimic the structure of reality, and which correspond to that part of the content of a scientific theory that is captured by its constituent general terms and their interrelations [3]
Reference Ontology	an **ontology** intended to provide an *informationally complete* representation of a domain
Application Ontology	an **ontology** representing the *portion of reality* which is relevant for some purpose in some community
Assessment Instrument Ontology	an **application ontology** describing the *portion of reality* covered by an **assessment instrument**
Data Collection Ontology	an **application ontology** describing the *portion of reality* covered in a **data collection**
Data Item	a **RA** that is intended to be a truthful statement about something (modulo, e.g., measurement precision or other systematic errors) and is constructed/acquired by a method which reliably tends to produce (approximately) truthful statements (modified from [4])
Measurement Datum	a **data item** that is a recording of the output of a measurement. [4]
Directive Information Entity	an **ICE** whose *concretizations* indicate to their *bearer* how to *realize* them in a *process* [4]
Conditional Specification	a **directive information entity** that specifies what should happen if a trigger condition is fulfilled [4]
Rule	an executable **conditional specification** which guides, defines, or restricts actions [4]
Bridging Axiom	a **rule** specifying how a **RA** should be interpreted in terms of an **application ontology**
Data Format Specification	the information content borne by the *document* published defining the specification (modified from [4])
Plan Specification	a **directive information entity** that when *concretized* is *realized* in a *process* in which the *bearer* tries to achieve the objectives, in part by taking the actions specified [4]
Assessment Instrument	a **plan specification** designed to compile **data collections** reliably, validly and reproducibly

Figure 4.9 Information Artifact Ontology Definitions (Ceusters, 2012)

For example, the IAO has its own information dictionary—a comprehensive hierarchical taxonomy of information types and definitions that describe the *aboutness* of information. In a paper entitled "An Information Artifact Ontology Perspective on Data Collections and Associated Representational Artifacts," Ceusters provides a mini dictionary of data definitions and refers to Aristotle as an influence. In the dictionary, various information types are referred to, with some of the highest classes including Information Content Entities (an entity that is generically dependent on some artifact and stands in relation of aboutness to some portion of reality), Representational Artifacts (an Information Content Entity which is believed to represent a portion of reality external to the representation), Representational Units (a Representational Artifact which according to the structural conventions it is designed, is not built out of any other RAs), Composite Representations (a Representational Artifact built out of constituent sub-representations as its parts), Data Items (a Representational Artifact that is intended to be a truthful statement about something (modulo, e.g., measurement precision or other systematic errors) and is constructed/acquired by a method which reliably tends to produce (approximately) truthful statements), and Directive Information Entities (an Information Content Entity whose concretizations indicate to their bearer how to realize them in a process). Lower subcategories descend from these top level categories.

As previously stated, the BFO is an upper-level ontology. Other upper-level ontologies have included Cyc (Cyc is one of the outdated ontologies referred to in chapter two). Smith told me that he did not want to give me his opinion on Cyc but he did say that ontologically speaking it does not satisfy what he sees as being minimal requirements. Still, Grenon (Smith’s coauthor on the SNAP and SPAN papers) came out of the Cyc world and he and Smith worked closely to create BFO and are co-authors on the initial papers in the BFO method. Since then, BFO has been co-managed and co-directed by Smith and a number of people. BFO has become very much a group exercise yet it is very hard to keep even a very small, upper level ontology in a good form—this is a hitherto unanticipated problem. BFO is small and changes slowly but to get good definitions of the terms and good axioms to make it work with the software involves all kinds of considerations which are difficult to resolve. It involves people with a number of different kinds of expertise in logic, in software, in programming, in social engagement with standards organizations, in applications of the ontology to building other ontologies which are domain specific rather than domain neutral, and so forth. The book (*Building Ontologies with Basic Formal Ontology*), Smith thinks, is a fairly successful representation of principles which they have developed over the years, tested over the years, and which are sound and which will survive. The difficulties – some of which have been covered in this dissertation – are not covered in the book. The version of BFO that exists now, BFO 2.0, is sound and stable, but it cannot be used in all the ways that people would want it to be used because the software is not amenable.

The success of ontology work is very much a difficult phenomenon that requires a large number of users. One ontologist explained to me that they think of the term "success" in ontology work in the following way. Suppose that you are building a telephone network for a country and you have really fantastic hardware and really fantastic cables, WiFi, or whatever it might be, and a really fantastic way of keeping track of telephone numbers, but there are not more than three subscribers. This would not be a successful telephone network. Success for ontology is, to a very large degree, a function of the numbers of users. BFO has been used by something like 130 different ontology groups. 100 or so of them are biomedical and there are a few groups in different areas who are applying BFO to topics such as financial services, developmental nanotechnology, or military projects. In that respect, BFO is by far the most successful ontology in that it has the greatest number of users. DOLCE did have, and still does have a number of groups who use it, but it is a relatively small number of groups compared to BFO.

The Ontology for Biomedical Investigations (OB) is a domain ontology that uses the BFO. If the BFO from a philosophical point of view is a bit like basic metaphysics then OB is the philosophy of science, particularly experimental science. OB is a BFO-based ontology which is designed to give the possibility of describing not the data created by an experiment but the processes which led to that data. The funding processes, the sampling processes, the staining processes, the measuring processes, the publishing processes, and so on. The point of this is that when dealing with complicated experimental results in medicine, or biology, you need to know how those results were acquired and processed. For that purpose, you need some kind of control vocabulary for describing the processes which led to those results.

Smith and Ceusters (2011) state that there are four main upper-level ontologies currently available today. These are the BFO, DOLCE, SUMO, and OpenCyc. They are, more or less, all strict upper level ontologies that do not contain representations of material entities—things like chemicals, rocks, and particles or other entities that would typically fall into specialized fields and discourses that are represented by domain ontology. BFO was purposefully built to be as tiny and abstract as possible so that it could succeed in the specific task of acting like the glue that holds domain ontologies together (Figure 4.9.1). In this way, the BFO can integrate many heterogeneous ontologies to create semantic interoperability, in the same way that those ontologies themselves create semantic interoperability among data. BFO operates at a higher dimension, as it were. DOLCE (Gangemi, Guarino, Masolo, Oltramari, & Schneider,

2002) has been very historically successful and has a high amount of users. DOLCE and BFO share many similar philosophical distinctions in their methodology, however DOLCE remains fundamentally different in that it focuses on “linguistic and cognitive engineering.” As such, DOLCE’s domain scope includes specific conceptual objects that would not be included in the BFO, such as fictional entities or things that can only be represented by concepts. There is an argument to be made that this makes DOLCE weaker from the perspective of scientific applied ontology-work which seeks to create the widest and tightest amount of semantic interoperability between ontologies.

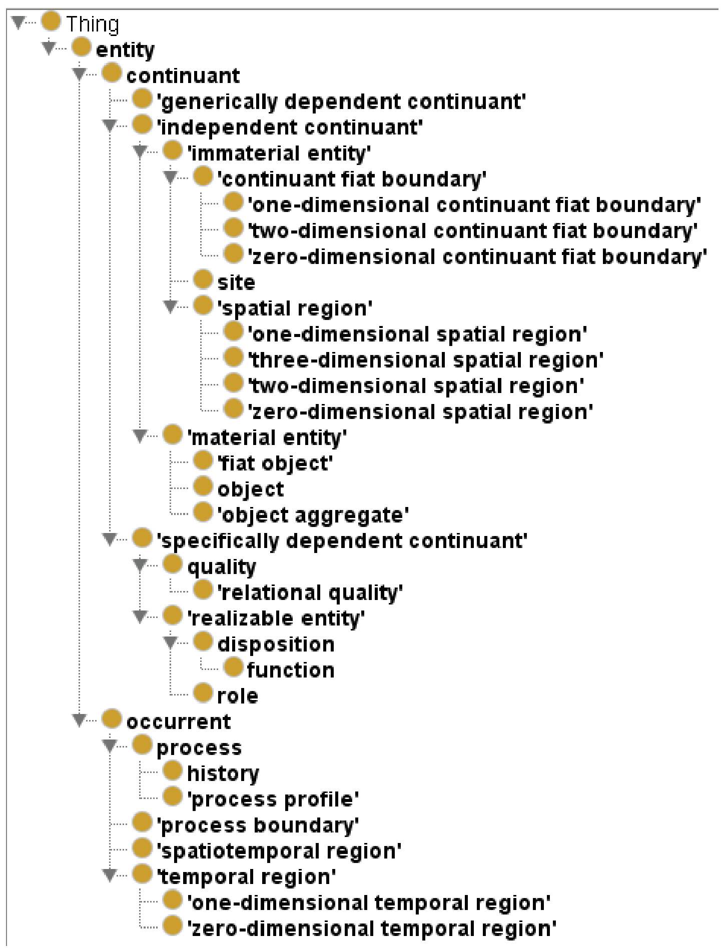


Figure 4.9.1 Basic Formal Ontology

SUMO has also proved historically valuable (Niles & Pease, 2001). However, as Smith and Ceusters note, the fact that SUMO "contains its own tiny biology ('protein', 'crustacean', 'body-covering', 'fruit-Or-vegetable') means that it cannot support the strategy of downward population that has proved so useful to scientists in the case of BFO, since biologists are unlikely to find SUMO's definitions (and selection) of biological terms acceptable" (p. 181). Remembering back to referent tracking and the notion that some biologists would rather share their toothbrush than a gene name, what this means is that it is unlikely that scientists would choose an upper-level ontology that attempts to create domain-specific names since those scientists likely would already be using their own domain-specific names and likely require the upper-level ontology for organization. Both DOLCE and SUMO are preferable to Cyc, however, since Cyc has ties to many common terms that are not used by scientists. Smith and Ceusters state that the primary problem with Cyc is that it does not "strive for consistency among the various 'microtheories' which form its parts" (p. 181). "Hence the very goal of creating a single consistent suite of interoperable ontologies which would capture the terminological content of biomedical science – which is from our point of view the only coherent strategy for achieving ontology-mediated data integration in the domain of the life sciences – is undermined by Cyc's own paraconsistent logical structure" (p. 181).

Smith and Ceusters embrace what they call the *Ontological traffic law principle*:

> *Ontological traffic law principle:* Ontological standards, including a common upper-level ontology and standards governing syntactical uniformity, are indispensable to every successful large-scale ontology development initiative, and this is so even if they are selected arbitrarily provided they enjoy widespread assent among those working in the relevant research community (p. 182).

An example of this would be the “law according to which all terms within an ontology should be nouns and noun-phrases that are singular in number” (p. 182). However, there are many scientists and engineers who do not agree with ontological realism or the BFO method. Many scientists and engineers still take sides in the realism versus conceptualism debate and there are a variety of blogs that discuss this issue. The blog OntoGeek (https://ontogeek.wordpress.com/) contains a series of posts on applied ontological realism with titles such as “Realism, Really?” and “Yes, really.” (OntoGeek doubled down on realism). Other blogs contain views against realism, such as those expressed in An Exercise in Irrelevance’s (http://www.russet.org.uk/blog/) post “Why Realism is Wrong.” There is also some tension between BFO and OWL. One ontologist related the following to me:

> I refuse to collaborate in any way on attempts to express ontologies by means of OWL. I won't do it and I don’t take any responsibility for what is there. An example is for instance the Ontology for General Medical Science (OGMS). I am one of the co-authors of the paper where the descriptions, the definitions, and the axioms were given for what became the OGMS. But I disagree with how OGMS is represented in OWL, simply because OWL is not expressive enough to really be able to express everything that is important if you want to have a good description of first order reality (J. Doe, phone interview, September 9, 2015).

OWL is used rather extensively in natural language processing where language is viewed as constructed, indicating that there are contexts in which OWL may be preferred to BFO. I asked the ontologist what is good enough and whether OWL was outdated. It is not outdated, he told me, because it is still extremely heavily used. However, OWL is built on the assumption, adopted by many computer scientists, that “bullshit in, bullshit out.” I assume that what the ontologist meant by this was that many technicians involved in ontology work do not see the semantic problem as being their problem and that they would rather produce ontologies that work and that can "consistently reason with bullshit." Such individuals call themselves ontologists, they said, but do not pay attention to semantics. According to the ontologist, the logic goes something like "Yeah well, but there is nothing better for the time being, and at least you can say some things in OWL, so then it is better to be able to say something than nothing at all." The ontologist says they are more principled in the sense that they refuses to work on a solution which cannot be perfect, though they perhaps underestimate the degree to which natural language processing communities deal with semantics in human language.

I then asked what areas they thought needed the most attention in ontology work. What areas need more development and research? The first thing, they said, is that there are different levels—the basic principle of referent tracking is that anything you want to say something about should be uniquely identified, but this principle is not widely adopted. It is adopted for a few things like patients who get unique identifiers, doctors, organizations, and x-ray machines, but not much else. The ontologist's claim is that everything that you want to say something about should first be given a unique identifier, but not before checking to see if someone else has already created one. If you are sure that nothing has been said about what you want to describe then you assign a new unique identifier. That is a principle that needs to be there. Secondly, the ontologist thinks that all descriptions should take serious consideration of the dimension of time more seriously.

> It is nonsense to come up with statements like "all humans have a heart," because that's not true during heart replacement surgery. Or that all humans have two legs. Look at the street. So those things should just not just be said like that. You should at least specify a certain type. You should say that whenever there is a human leg, it has been part of some human being at some time, but it might have

> been cut off. So those kind of statements, I think that's important. Assertions, which, both in ontologies as in individual statements, take the time elements seriously (J. Doe, phone interview, September 9, 2015).

Other ontologists are less critical of OWL. He explained to me that first order logic is not from a computational point of view ideal. If you have an ontology which is using first order logic, it will very often not execute queries. The computer just will not be able to result queries in a reliable way. The computer will not be able to reliably execute queries, not because the computer will create false results but because the computer will never end its process of trying to work out what the results might be—it will take an infinite time, which is not good. In consequence to this the military and other agencies have been working on trying to find fragments of first order logic which will execute reliably. The current preferred fragment of first order logic is OWL, one of whose founders is Tim Berners-Lee. Lee founded OWL as part of what is called the Semantic Web, which was an idea to create a version of the Web which would enable the kind of search, combination, comparison, and reasoning that has been covered in this dissertation. OWL is today the default language of choice for ontology work. OWL's expressivity is weak but its computable properties are good. One ontologist told me that there are features of BFO which they would like to be able to express using OWL which they cannot. There are some things that can be expressed easily using first order logic which cannot be expressed easily using OWL. The tradeoff is that researchers involved in the BFO do not want to create a BFO which no one will use because it is too complex and so they are trying to find a way of resolving this issue. Keeping BFO simple, keeping its capacity to deal with time, which is where the problems primarily arise, but without sacrificing the computational qualities that OWL provides.

Smith is now attempting to make people take the view that BFO exists in different forms. There is an OWL version, there is a first order logic version, and there is also an English version. All of these versions have their good qualities and their bad qualities, and all of them should be kept alive, and will be kept alive, one ontologist said. They should be kept alive in such a way that the three versions are compatible as far as they can be. But because OWL has weak expressivity, there will be parts of BFO that cannot be expressed in OWL. OWL is like the alphabet (A, B, C, D…), while BFO is like *and* or *not*, *thing*, *process*, *adjective*. BFO is the next level of structure after the alphabet. There is some structure in OWL but it is not a competitor to BFO. "Some people think that if you have an artifact which has been built using OWL then you have an ontology," the ontologist said. "I think that in order to have an ontology which is useful for anything you need a lot more than just a piece of correct OWL code." The ontologist is partly right—in the empirical natural sciences BFO is the more desirable of the two ontologies since it follows the rules of ontological realism. However, there are natural language processing communities that continue to find OWL useful in basic research in that domain.

4.6 The Open Biological and Biomedical Ontologies Foundry

The Open Biological and Biomedical Ontologies Foundry (OBO Foundry) (http://obofoundry.org/) is a coordinated attempt to evolve ontologies to support biomedical data integration. Currently, the OBO Foundry is one of the largest coordinated grouping of ontologies in the world (Figure 4.9.2). As noted in chapters one through three, the value of data is increased when organized in a form that allows for them to be semantically interoperable. Applied ontologies are an attempt at producing such semantic interoperability, yet the profusion of a great many ontologies replicates the same problem that exists when attempting to combine different datasets. Many ontologies in one area of research means that those ontologies carry their own meanings, definitions, and relations and consequently cannot be combined with other ontologies. The Foundry is a strategy to remedy this problem by employing standards and editors (the OBO Foundry Operations Committee) to make sure that semantic interoperability can be maintained across domain ontologies. The Foundry follows BFO principles and Smith acts as one of its coordinating editors. The Foundry also includes a list of principles that are intended as normative constraints according to which submitted ontologies for review will be evaluated. These principles are considered best practices that ontologists should follow even if they are not planning on submitting their ontology to the Foundry. The principles listed on the OBO Foundry website are divided between MUST and SHOULD requirements. To be added to the Foundry, ontologies must be openly available, created using a common formal language, include metadata for describing changes in the ontology, use coherent natural language definitions of top-level terms, include textual definitions, use unambiguously defined relations (this is a controversial rule in the OBO Foundry), provide documentation, document that the ontology has multiple users, be carried out in a collaborative fashion, have a contact person, follow naming conventions listed in Schober et al. (2009), and be maintained in light of scientific advancement.

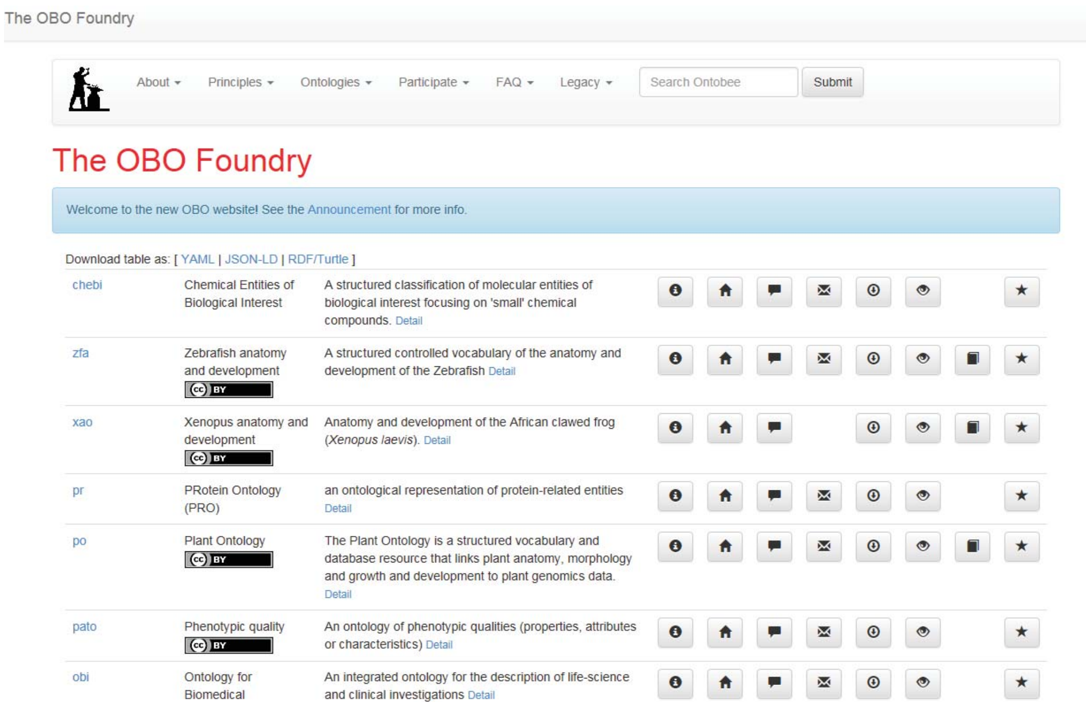


Figure 4.9.2 The OBO Foundry

One ontologist explained to me that if someone wants to get a label from the OBO Foundry they need to pass through a panel of experts, including the OBO Foundry Operations Committee, who will look at the documentation on the ontology and the principles that are used and make a decision on whether or not the ontology deserves the OBO Foundry's stamp of approval to be included in the group. Unlike the Foundry, a similar attempt at collecting ontologies – the NCBO's BioPortal (http://bioportal.bioontology.org/) – does not enforce such standards and allows for the inclusion of many different ontologies that are not semantically interoperable. Of the BioPortal, the ontologist told me:

> That's the biggest collection of ontology junk, because there is no principled approach at all in what is submitted there. The majority of those ontologies even violate the semantics of the description language used. The BioPortal is the place where you can get all the biomedical ontologies without any quality criteria. The only thing that they do is to say there are so many classes and there are so many

> individuals and there are so many relations and that's it (J. Doe, phone interview, September 9, 2015).

There is, however, disagreement among the ranks of the OBO Foundry as well and this has been well documented on the OBO Discuss mailing list (http://obo-discuss.2851485.n2.nabble.com/). There are currently debates about clarifying the wording of certain principles and expanding the criteria used to review the ontologies for consideration in the Foundry, indicating that the Foundry itself is not immune to the type of controversy that affects entities such as the BioPortal. Not all members of the Foundry community are happy with the standards as they currently exist, indicating that there are problems even though such principles are in place. Similar debates have existed for years on the mailing list. Below is an example of Smith responding to a comment left by Peter Karp on the OBO Discuss mailing list on July 16th 2010 (Karp's text has arrows while Smith's answers do not and are located in between Karp's text):

```
At 12:52 PM 7/16/2010, Peter Karp wrote:
>My take on this discussion is
>
>(a) The paper by Dumontier and Hoehndorf makes a number of excellent
>points. It will of course be interesting to see what Barry et al
>have to say in response.
>
>(b) The OBO Foundry Principles are so unclearly expressed that they
>are not deserving of all the key pokes they have generated. The
>instantiability principle is one example of such a vaguely formulated
>principle (in fact, it is rather ironic that a group concerned with
>ontologies has formulated such unclearly stated principles for
ontologies).
>
>Most of the principles on that page require further discussion,
>explanation, and justification, and I recommend that the OBO Foundry
>group create an accompanying web page that provides a more in-depth
>discussion of each principle.

I believe that we have been trying to do this, in very many papers,
lectures. videos, etc. But we can of course try harder.

>  Examples would be particularly helpful --
>"near misses" are always a helpful way of clarifying definitions.
>For example, it would be helpful to know what Barry or others think
```

>are possible ontology terms that people might want to define that
>should not be defined because they do not correspond to instances
>in reality.

unicorn
mermaid
leprechaun
absent nipple
absent leg
cancelled performance
entity that is simultaneously an instance of Heart failure, Tooth decay, and Pregnancy
single-celled mammalian organism

> What does it even mean to correspond to an instance in
>reality?

An instance is something that exists in space and time. Often instances can be observed or measured. For a referring expression to correspond to an instance means nothing more sophisticated than for it to refer to or name that instance. E.g. 'Peter Karp' corresponds to Peter Karp. Part of the problem is that we are dealing with issues so basic that there is very little further that we can say that would illuminatingly explain them.

> Is a regulation event an instance in reality?

Yes

> Is an
>experiment plan an instance in reality?

Yes

> Is the process of transcription
>an instance in reality?

Any given transcription process occurring in some given place and time is an instance in reality; an instance of the type transcription process?

> Or are these entities simply cognitive
>constructs?

If they exist independently of scientists' descriptions of them then they are not cognitive constructs.

>(No doubt these issues are discussed in Barry's
>papers, but the OBO Foundry principles should be comprehensible
>on their own.)

We are, indeed, doing our best.
BS

>P

In the above exchange, Karp writes that the "OBO Foundry Principles are so unclearly expressed that they are not deserving" of their success and that the principles are "vaguely formulated." Karp notes that it is "ironic that a group concerned with ontologies has formulated such unclearly stated principles for ontologies," further stating that "it would be helpful to know what Barry or others think are possible ontology terms that people might want to define that should not be defined because they do not correspond to instances in reality." Smith (Barry) replies that they are attempting to clarify the principles and offers Karp a list of "near misses" that would not qualify as entities in the ontology: unicorn, mermaid, leprechaun, absent nipple, absent leg, cancelled performance, entity that is simultaneously an instance of heart failure, tooth decay, and pregnancy, and single-celled mammalian organism. Perhaps most interestingly, when Karp asks "What does it even mean to correspond to an instance in reality?" Smith simply answers back "An instance is something that exists in space and time." This exchange represents a small portion of a larger discussion around fundamental issues within the BFO and OBO Foundry communities. It is interesting to see such debates continuing to exist around some of the common BFO principles when applied to entities such as the Foundry. Such debates are indicative that even the method of ontological realism is not devoid of debate and controversy, that rhetorical appeals must occasionally be made to convince other members of the BFO and Foundry communities, and that there are applications of the ontological realism methodology that can potentially lead to perceived problems. The BFO Discuss group and the OBO Discuss mailing list contain numerous examples of such controversies.

CHAPTER 5. CRITICAL DATA STUDIES, ONTOLOGY, AND ETHICS

5.1 Chapter Summary

This chapter discusses ethical concerns that may arise out of the application of computational ontology to the social sphere, introduces the field of CDS, and explains how CDS is concerned with ethical and epistemological issues related to data (boyd & Crawford, 2012; Dalton & Thatcher, 2014; Kitchin & Lauriault, 2014). It describes some of the concepts that inform CDS as a framework, including the concepts of data assemblages (Kitchin, 2014) and infrastructural inversion (Bowker & Star, 2000). Following work on ethical reasoning in Big Data (Steinmann, Matei, & Collmann, 2016), I present a heuristic approach to ethical reasoning in computational ontology and the data that constitute it that is built on the principles of autonomy, beneficence, nonmaleficence, and justice (Beauchamp & Childress, 1979/2013). It proceeds by offering a critique of BFO principles and methods when applied to contexts outside of natural scientific research, including applied social ontology in the form of the Military Ontology. The Military Ontology is supported by the BFO and the project was a direct response to a Chairman of the Joint Chiefs of Staff Instruction (CJCSI) entitled “Horizontal Integration (HI) of Warfighter Intelligence.” The chapter argues that the subject of CDS should include not only things like social media and the internet but also emerging media like the BFO and other technologies of individuation that affect social ontology. CDS focuses on the techniques, technologies, institutions, and methods involved in the production of data. The chapter explores ethical reasoning in applied ontology work and ends by discussing virtual science.

5.2 Critical Data Studies

CDS is a subfield of STS and a new approach to studying data and their infrastructures, one that may be considered a subfield of Science and Technology Studies. Technologies of individuation that process informational ontology are the general subjects of CDS, which is a theoretical approach (boyd & Crawford, 2012; Dalton & Thatcher, 2014; Kitchin & Lauriault, 2014) that questions data structures according to a mixture of ethical, critical, organizational, practical, and policy-oriented criteria (Iliadis & Russo, forthcoming). CDS is a growing field of research that focuses on the unique ethical and epistemological challenges posed by data infrastructures. Articles in CDS have shown that data are never 'raw' and that they must constantly be negotiated, maintained, and interpreted (Gitelman, 2013). Rather than treat data infrastructures as only scientifically empirical and therefore largely neutral phenomena, CDS advocates the view that data should be seen as always-already constituted within wider data assemblages (Kitchin, 2014). Assemblages is a concept that helps capture the multitude of ways that already-composed data structures inflect and interact with society, its organization and functioning, and the resulting impact on individuals' daily lives. CDS questions the many assumptions about data infrastructures that permeate contemporary literature on ICTs and society by locating instances where data may be naively taken to denote objective and transparent informational entities to expose their communicative power.

Computational ontology is not immune to ethical questions and problems from a multitude of perspective. For example, one ontologist said that there are powerful incentives in the software industries to invent something new for each body of data and each new customer. Some of these incentives relate to genuine issues of data privacy and security. Often, however, the resultant siloing of data is indefensible on ethical grounds. The idea here is that it is wasteful and counterproductive to produce more and more heterogeneous data when one of the important objectives in ontology work currently is to produce semantic interoperability. Another data scientist put it to me that there are simply too many databases and that there should be a tax for each new one that is created (a database tax). I am more interested here, however, in some of the ethical implications of ontologies when applied to specific domains, particularly when they are applied to areas that exist outside of the natural sciences. How should issues relating to causality, quality, security, and uncertainty in computational ontologies be taken up when these technologies of individuation are applied to domains in the social, for instance? Such questions are difficult to formulate and require a set of conceptual tools to help set up a framework that can enabler ethical investigative inquiries into applied computational ontology and its ethical consequences. Ontologies are a form of emerging media (Smith, 2016) in two sense in that they are, first, a new form of technology that has been developed to mediate data and transform it among computers, networks, and between individuals and institutions, not to mention methods and standards, but also, second, emerging in the sense that the data that they produce are themselves emergent—the partitions that are created by computational ontologies individuate entities and are thus productive in re-ontologizing the world. Technologies of individuation are theoretically important for understanding how computational ontologies and similar forms of emerging media manipulate reality according to an informational ontology that partititons reality according to levels of granularity and informal abstraction. As technologies of individuation, computational ontologies output emergent universal data structures that appear only after a process of informational individuation has occurred which creates data entities that did not exist previously. Such emerging forms of media require careful inspection and consideration of their application to the social world, among others.

To understand computational ontologies as technologies of individuation, emphasis should be place on their data assemblages (Kitchin, 2014). The apparatus and elements that make up a data assemblage include systems of thought, forms of knowledge, finance, political economy, governmentalities and legalities, materialities and infrastructures, practices, organizations and institutions, subjectivities and communities, places, and the marketplace where data are constituted (Figure 5.1). *Assemblage* is a term that comes from the French philosopher Gilles Deleuze and should be understood here as a structure that emerges from a variety of social entities existing at multiple scales (local, national, international) that exert power. A data assemblage is a powerful complex of entities that form the underlying production of data at multiple levels and in a plurality of domains.

Apparatus	Elements
Systems of thought	Modes of thinking, philosophies, theories, models, ideologies, rationalities, etc.
Forms of knowledge	Research texts, manuals, magazines, websites, experience, word of mouth, chat forums, etc.
Finance	Business models, investment, venture capital, grants, philanthropy, profit, etc.
Political economy	Policy, tax regimes, public and political opinion, ethical considerations, etc.
Governmentalities and legalities	Data standards, file formats, system requirements, protocols, regulations, laws, licensing, intellectual property regimes, etc.
Materialities and infrastructures	Paper/pens, computers, digital devices, sensors, scanners, databases, networks, servers, etc.
Practices	Techniques, ways of doing, learned behaviours, scientific conventions, etc.
Organisations and institutions	Archives, corporations, consultants, manufacturers, retailers, government agencies, universities, conferences, clubs and societies, committees and boards, communities of practice, etc.
Subjectivities and communities	Of data producers, curators, managers, analysts, scientists, politicians, users, citizens, etc.
Places	Labs, offices, field sites, data centres, server farms, business parks, etc., and their agglomerations.
Marketplace	For data, its derivatives (e.g., text, tables, graphs, maps), analysts, analytic software, interpretations, etc.

Figure 5.1 Apparatus and Elements of a Data Assemblage (Kitchin, 2014)

Data assemblages “frame how data are produced and to what ends they are employed” (Kitchin, 2014, p. xvi). Data can be conceived as “the central concern of a complex sociotechnical assemblage” that is “composed of many apparatuses and elements that are thoroughly entwined, and develop and mutate over time and space. Each apparatus and their elements frame what is possible, desirable and expected of data. Moreover, they interact with and shape each other through a contingent and complex web of multifaceted relations” (Kitchin, 2013, p. 24). Computational ontology as a technology of individuation should similarly be approached as consisting of these various categories of data assemblage. A critical approach to computational ontologies and their data should focus not only on the data that are generated by the ontology but also the institutions, financing, and forms of knowledge that contribute to the technology of individuation so as to expose the degree to which such technologies are themselves productive of data which are never neutral. Semantic data in particular as the main aim of computational ontology should be considered in light of the apparatus and elements presented here. Such individuating technologies should be critically analyzed by focusing on each apparatus that functions as a condition of possibility for the data assemblage. Computational ontologies in particular are suited to such studies since they are comprised of multiple histories, groups, institutions, standards, and methods that are often ignored in favor of scientific progress.

Infrastructural inversion (Bowker & Star, 2000) is the practice of foregrounding the infrastructures that often remain hidden in scientific research (and ethnographic research on scientists). A CDS approach should pay attention to such infrastructures and foreground them rather than the content that is produced by technologies. CDS should focus in on those structures and infrastructures where data are nested and where they emerge as individualized entities. Data themselves sit only at the surface level of the individuating structures and technologies of individuation that are formative of data and produce stable entities. Informational ontology is another way of thinking about CDS—critically analyzing the world perceived as consisting of layers of informational abstraction. The goal of CDS should be in actualizing or laying bare some part or heretofore unforeseen section of informational ontology to expose a new reality. Infrastructural inversion "is a struggle against the tendency of infrastructure to disappear" (Bowker & Star, 2000, p. 34). Bowker and Star write that

> Infrastructural inversion means recognizing the depths of interdependence of technical networks and standards, on the one hand, and the real work of politics and knowledge production8 on the other. It foregrounds these normally invisible Lilliputian threads and furthermore gives them causal prominence in many areas usually attributed to heroic actors, social movements, or cultural mores (2002, p. 34).

Methodological themes for infrastructural inversion include ubiquity (how classification schemes saturate an environment), materiality and texture (standards are material), the indeterminacy of the past (there are multiple times and multiple voices), and practical politics (of classifying and standardizing). Bowker and Star's method of infrastructural inversion combined with Kitchin's approach to data assemblages provides enough conceptual framework to enable a critical data study into complex emerging technologies of individuation like computational ontologies. Infrastructural inversion can also be a generative resource in digital scholarship (Kaltenbrunner, 2015). The concept of infrastructural inversion implies that "infrastructure is not a specific thing such as tubes and wires, but a relational state that obtains when actors working in different parts of a historically grown, cooperative work setting achieve a smooth coordination of their individual activities" (Kaltenbrunner, 2015, p. 4). Technologies of individuation like computational ontologies are coordinated by not only physical technologies but also digital technologies such as email lists, blogs, and message boards. Such organizing projects require a deep CDS analysis of digital environments to understand their power and the collective way in which such technologies are built and used.

CDS should specifically provide communication with the following. First, CDS offers a new ethical methodology from which to conduct inquiries related to ICTs as an empirical endeavor. An individuative methodology would seek to proceed by articulating instances of the modulation of communicative processes themselves, rather than in the simple transmission of meaning or data between pre-given, already individuated entities. For example, whether studying empirical evidence in doctor-patient health communication or the analysis of vast quantities of data in social network analysis, an individuative methodology would seek to measure, uncover or understand those communicative structures that modulate in the act of communication and that perpetuate by virtue of an individuative flexibility. What variable characteristics of the formal consultation setting are responsible for trends that develop in interpersonal communication? How do reflective properties inherent in the visibility of a wiki edit history potentially alter future edits? These are the structural qualities of modulation that an individuative methodology would seek to uncover. Second, CDS offers a new conceptual toolbox and specialized terminology with which to frame future discussions on entirely new communicative phenomena: the language of technologies of individuation. Instances of modification in the technical evolution of objects such as ontologies, programs, and games or their material effects on social reality can be referred to as points of individuation. Moments where once-separate levels of communicative or informational properties are linked and give way to something new can be referred to as acts of disparation. Third, CDS bypasses a longstanding debate in STS, one that affects the future of communication studies also. A CDS view of informational ontology finally puts aside the social construction vs. technological determinism deadlock in STS and instead considers the human that is present in the informational ontology as an assemblage or ensemble. Communication research into interfaces and human-computer interaction stand to benefit from a CDS approach to technology and embodied interaction where the point is less about the separation of the human from the technical than it is about the successful interoperability of the ensemble. Fourth, CDS shifts the discussion from paradigms of closed ecologies to wide-open informational paradigms—informational ontology as a form of ISR can contribute to communication studies not by offering predetermined boundaries of inquiry as in ecology but by recommending an open informational realism that is amenable to radical interdisciplinary forms of research. This dissertation accomplishes these four goals by treating applied computational ontologies (the BFO) as technologies of individuation and by approaching them from an ethical framework that seeks to expose the data assemblages that form their ability to modulate reality.

CDS must also be done within a particular context. Following Nissenbaum (2011), CDS must study data in social, government, commerce, science, and other contexts. Of context, Nissenbaum (2011) writes

> Contexts are structured social settings characterized by canonical activities, roles, relationships, power structures, norms (or rules), and internal values (goals, ends, purposes). […] [S]ocial life comprises structured, differentiated spheres, whether labeled and theorized as "fields," "institutions," "structured social (p. 132).

Contexts play an important role in the way that we understand privacy but also in the way emerging technologies are nested in the social world and situated among actors in the public. Computational ontologies individuate entities and structures as they re-ontologize the world, thus they are formative of contexts at the same time that they are embedded in them, making them unique in the realm of emerging media. When studying computational ontologies emphasis should be placed on the contexts in which they are applied.

Lastly, CDA needs a framework or rubric to study data and its impact. Steinmann, Matei, and Collmann (2016) offer a matrix for ethical reasoning in Big Data that can be adapted to research on CDS, ontologies, and their data. Their work combines the ethical principles of Beauchamp and Childress (1979/2013) with the contextual approach propagated by Nissenbaum (2010), creating a principle-context matrix. The principles of Beauchamp and Childress were first formally presented in the context of biomedical ethics, however recently there has been an increase in the application of the principle to other fields such as engineering. Following the work of Steinmann, Matei, and Collmann (2016) and their adoption of Nissenbaum's (2011) context-dependent approach in the ethical reasoning in Big Data, I argue that ethics in CDS should similarly be analyzed in context-specific cases. This matrix can be used similarly to study applied ontologies and their ethical impact on agents and their environments in different contexts.

Table 1 Ethics Matrix (Steinmann, Matei, and Collmann, 2016)

	Social	Government	Commerce	Science
Nonmaleficence				
Beneficence				
Justice				
Autonomy				
Trust				

The ethical matrix is useful in considering computational ontology work since ontologies are applied in a wide variety of contexts, including social, government, commerce, and science contexts. However, there are many circumstances where multiple contexts can exist at once while overlapping each other. For example, in a military context, there may be elements of the social, government, commerce, and science overlapping all at once. While the ethical matrix does not suggest that each context exists at the exclusion of the others, emphasis should be placed on the idea that each context, like each principle, can be present at once. Computational ontologies as technologies of individuation span several contexts. Rather than being seen as associated only with science and realism, computational ontologies must be opened up to the many other domains of application in which they are found, including social, government, and commerce contexts, but also in specific fields such as natural language processing.

5.3 Ontology in Context and Social Ontology

I discussed ethics in applied ontology with Peter Yim and we came to the conclusion that ethics in applied computational ontology work can be approached at two levels. At the technical level, if one were to look at an ontology as being the representation of the shared understanding of the meaning of a thing, among members of a given community, in a way that can be processed by both humans and machines, then ontologies can be viewed as ethically neutral. The question is analogous to asking: "Are there ethical problems in applying logic?" or even "Are there ethical problems in applying statistics?" Yet, on another level, like all applications of technology made by humans, there is almost bound to be ethical issues—the more powerful the technology, the more issues will likely exist and need to be addressed. This will be analogous to asking: "Are there ethical problems in applying laser technology?" or "Are there ethical problems in applying bio-technology?" or even "Are there ethical problems in developing

Artificial Intelligence?" The key word is "applying"—what is the application? Who is applying? How is it applied? etc.

Computational ontologies can be applied in many domains outside of the natural sciences, including social domains. We have seen that in the context of natural science the method of ontological realism is ideal for success in applied ontology work—such ontologies depend on the realism of universal scientific laws and processes. Yet, the method of ontological realism embodied in the BFO has been applied in areas outside of the natural sciences, in fields such as military, intelligence, social planning, and economics, where arguments appealing to the universality of ontological realism become harder to defend. At the time of writing, BFO is currently applied to roughly 166 ontologies. Many of these ontologies are in fields outside of the natural sciences. I created a list of the ontologies that use the BFO, looked up their purposes and definitions, and applied them to the contexts provided by Steinmann, Matei, and Collmann's (2016) ethical matrix. Out of the 166 ontologies using BFO and its underlying ontological realism-based principles, 27 are or can be applied to a social context, 22 are or can be applied to a government context, 22 are or can be applied to a commerce context, and 155 are or can be applied to a science context (the vast majority in the natural biological and chemical sciences). Many of these contexts overlapped, and if the ontology was related to a technical method then I marked it as applied to all contexts. I did not attempt to apply the principles (autonomy, beneficence, nonmaleficence, justice) to these ontologies. Yet, one can see by the Bank Ontology, the Military Ontology, the Ontology of Social Participation, and the Spatiotemporal Ontology for the Administrative Units of Switzerland that there are ontologies that follow BFO principles that exist outside of the domain of natural science and more in the social or government contexts. In such contexts, BFO might encounter some ethical difficulties that are not present in the natural science ontologies. What I mean by this is that if the main benefit of the BFO over other ontologies is supposed to be (and indeed likely is given the astounding success of the BFO method) its commitment to scientific realism in the sense of adhering to the most accurate universal laws, facts, and patterns identified by science, then the application of this ontology in fields that are not explicitly and uniquely scientific may encounter problems.

A number of the ontologies that use BFO fall in the category of what has been referred to by the philosopher John Searle as social ontology (Searle, 2006). Searle first began thinking about social ontology in *The Construction of Social Reality* (1995) where he sought to provide the invisible structure of social reality—things like institutional facts, social phenomena, and social facts. Searle writes that social ontology "is both created by human actions and attitudes but at the same time has an epistemically objective existence and is part of the natural world" (2006, p. 12). The fundamental concepts that are needed to explain the existence of social ontology include: the distinction between observer-relative and observer-independent phenomena, the distinction between the epistemic and the ontological senses of the objective–subjective distinction, the notions of collective intentionality, the assignment of function, and constitutive rules. Perhaps the most important notion is that of status functions in institutional ontology, which Searle describes as "the glue that holds society together because they create deontic powers, powers that work by creating desire-independent reasons for action. Thus, social ontology locks into human rationality" (2006, p. 12). If the ontological realism endorsed by Smith and the BFO is dependent on the universal laws and structures of science, social ontology is the dependent on the invisible rules and laws that society follows—the former individuates scientific entities (natural kinds) while the latter individuates social entities (social kinds).

Smith is an adherent of Seale's social ontology and has continued the practice, developing and extending Searle's work through his own ontological work on the aforementioned document acts (2014), which take a social ontology perspective on things like institutional systems to which documents belong, positional roles within such systems, the production of documents, the connection of documents to reality through things like evidence, authentication and security documents, etc. The entities represented by social ontology (and its kinds that are represented in subfields such as document acts) are certainly real in the sense that they carry purpose in the world and exist as actual entities (there are things like presidents, mothers, deeds, decrees, documents, and so forth). Social kinds are said to be different from the natural kinds produced by science in that they are subjective and depend on mental attitudes (Searle, 1995)—social kinds are interactive and are malleable. Yet, social ontology as practiced in things like the Military Ontology depend on the ontological realism presented in the BFO while identifying entities that belong in the realm of social ontology. The methodological realism of the BFO and its relations might need to be ethically analyzed when applied to the types of data that belong to the realm of social kinds. A femur *is_a* bone—such statements should be uncontroversial in an ontology using the BFO in that they are universally true and not contested. But ال ناصر ع بد (Abdul-Nasir) *is_a* leader is obviously not the same type of statement in that it is not universally true and is the product of collective intentionality (of the military) and fiat. Such social ontological commitments should be subject to ethical review when expressed in computational ontologies. The Personal Name Ontology is used by the Military Ontology and is an example of social ontology (Figure 5.2).

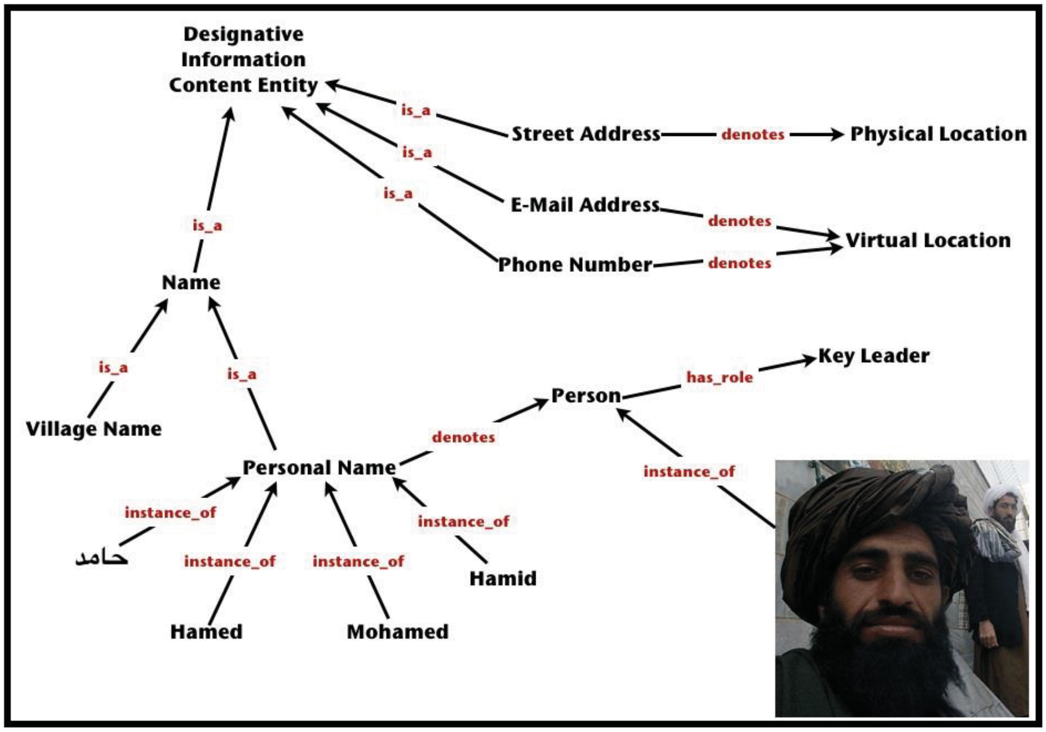


Figure 5.2 Personal Name Ontology

Does the ontology make sound and ethical definitions in terms of its entities and relations, has the data that will be processed by the ontology been ethically gathered, how do the well-honed principles of autonomy, beneficence, maleficence, and justice apply to the computational ontology and data? Some social kinds might be understandably taken as permanent—*father*, *child*, *group*, etc. Others, however, are born of fiat—things like *money* and *terrorist* are whatever hold the conditions that we ascribe to such entities. Computational ontologies should be subject to ethical review when the individuations that they create are related to social or other contexts outside of the natural sciences (though science itself should not be immune from similar ethical commitments). When ال ناصر ع بد (Abdul-Nasir) is individuated as a leader in the Military Ontology using BFO principles, there should be some extra ethical step before processing the name as a leader in an ontology that may produce any number of other (potentially negative) consequences based on that action. Similarly, something as seemingly benign as the Email Ontology should be reviewed according to the ethical ontology matric with principles to see if the potential individuations that may be produced in the ontology will lead to potentially negative consequences.

5.4 Military Ontology

Military ontology has been practiced by armies since the beginning of organized warfare. In *On War* (1832), Carl von Clausewitz begins by describing the mereology of war and the need for situational awareness and understanding:

> I propose to consider first the various elements of the subject, next its various parts or sections, and finally the whole in its internal structure. In other words, I shall proceed from the simple to the complex. But in war more than in any other subject we must begin by looking at the nature of the whole; for here more than elsewhere the part and the whole must always be thought of together (1989, p. 75).

Similar to how the OBO Foundry operates in the domains of biology, medicine, and health sciences, BFO principles are being put to use to create similar globally unified ontology standards in intelligence for entities such as the United States Military, the Central Intelligence Agency, and other intelligence agencies (Figure 5.3). I spoke to one ontologist about their work with the military and their involvement in military intelligence and work with the United Nations on development ontologies. “The area where I’m probably being most influential now is military intelligence,” they said. The ontologist worked on projects for the Intelligence and Information Warfare Directorate (I2WD) which consisted of an attempt to address the following problem: There is a large amount of data, for instance data pertaining to terrorists or something similar. This data comes from many different kinds of sources. Some of it is human intelligence data, some of it is signals intelligence data, and so forth. The intelligence analysts have to use this data to answer questions and there are various rules they use to get answers to questions. Typically, each intelligence analyst will understand the structures of the databases that he or she has to work with only superficially, except for a small number of databases, where they are experts. The intelligence analysts thus have good access only to a small fraction of the entire data available and do not have a quick way of gaining access to all the rest. This is a simple kind of problem and the reason for it was because the data in all of those databases was described in different ways.

Figure 5.3 Military Ontology Website Information about Basic Formal Ontology

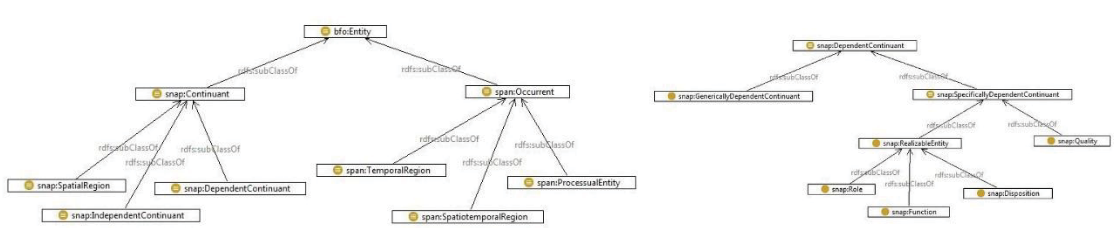
There was a lot of data about persons and in some of the databases the data about persons was organized with a column headed *person*. Then, under that heading, you would have the names of the person, in Arabic, for example. In other databases, the person data would be organized using a word like *human*, or *human being*, or *target of interest*, or *P*, or some acronym, which some people knew the meaning of, but some people did not. Which meant that there was no way of creating a single index of all the persons about whom data was available. Similarly, there was no way of creating a single index of all the data about places that people had data about, or about meetings, or about explosions. The data were created, like databases typically are, by using incompatible data labels. This tradition in database design has to do with the fact that you use acronyms because you want to squeeze as much data onto a single screen as you can. But every database engineer creates his or her own acronyms. The ontologist suggested that they should create a simple ontology and describe all the person-related acronyms with the single word, *Person*, and all the place-related acronyms with the word *Place*, and so on. It turned out to be a very useful thing—amazingly, useful. It changed the way the intelligence analysts worked, the ontologist was told, as they were not physically present when this happened. But they were told that it changed the way the intelligence community were able to use the data. Overnight they came back and said they wanted more ways in which they could use ontologies. This was a very simple but also a very successful first step within that particular community.

The Military Ontology website uses language that is very similar to ontological realism and describes military ontology as "devoted precisely to the representation of entities as they exist in reality" and as "a shared resource for disparate communities to communicate with each other." The website further states that the Military Ontology contains extensive material on ontology, BFO, and ontological realism and states that the purpose of the site is to

> Facilitate the rapid exploitation of the methods of ontology by the U.S. Military, for the purposes of: Providing an ontological resource to support representing, defining, and relating constituent elements in various military domains; identifying best practices in ontology development and creating a repeatable process; creating interoperable and consistent semantics in a modular fashion; facilitating the Horizontal Integration of Warfighter Information.

The Horizontal Integration of Warfighter Information is contained in a paper with the same title written by Smith and several coauthors (Smith is the first author) which was first presented at the Seventh International Conference on Semantic Technologies for Intelligence, Defense, and Security October 23-15 2012 at George Mason University in Fairfax, Virginia. The published conference proceedings describe "a strategy that is being used for the horizontal integration of warfighter intelligence data" which "rests on the development of a set of ontologies" that are meant to bring about semantic interoperability. The project was a direct response to a Chairman of the Joint Chiefs of Staff Instruction (CJCSI) entitled "Horizontal Integration (HI) of Warfighter Intelligence." The CJCSI states that HI is the "set of processes and capabilities to acquire, synchronize, correlate, and deliver national security community data with responsiveness to ensure success across all policy and operational missions." The Military Ontology is supported by the National Center for Ontological Research at the University at Buffalo and Smith's BFO has become the main ontology in use by the military to answer the question posed by the CJCI document.

Like the example of Bob's hyperthyroidism which was described at the beginning of this dissertation, the Military Ontology states that the "fundamental problem is that while there is only a single reality, stovepipe systems represent this single reality as if there were multiple realities by using disparate terminology and process descriptions." Unlike ontological realism in the natural sciences however, the Military Ontology is applied to data that represent a form of social ontology which are not quite the same as scientific ontology. For example, the Military Ontology uses BFO terms like independent continuent, dependent continuent, and processual entity and contains adapted SNAP and SPAN images on its webpage (Figure 5.4).



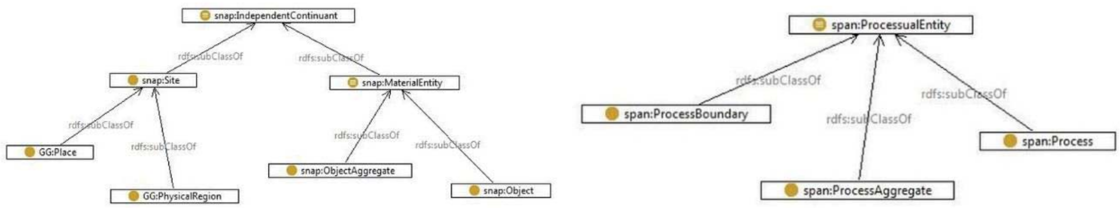


Figure 5.4 SNAP and SPAN in the Military Ontology

The Military Ontology seeks to apply BFO ontology work in multiple areas that exist outside of the domain of science, including the Ontology of Counterinsurgency, which is described as a “stratified phenomenon consisting of physical and cognitive strata” (Mandrick, 2008). Counterinsurgency is waged for “control of the cognitive stratum” and “to influence the population to reject the insurgent movement and to gain their support of the government” which requires “focusing upon the collective psychology of the population in order to develop the correct strategy” (Mandrick, 2008). Now, while there can certainly be SNAP and SPAN entities in an ontology like the Ontology of Counterinsurgency, the data in the ontology will be about entities that are defined using a social ontology rather than a scientific one. This is not wrong in itself, however such entities need to be thoroughly reviewed before being included in an ontology that uses the BFO principles, if the BFO principles are said to rest on the universality of scientific evidence—it is difficult to see how terms like *insurgent* are purely scientific.

There are other ontology projects in the military that are similarly based on a combination of BFO and social ontology—many of these are discussed at Semantic Technology for Intelligence, Defense, and Security which is a series of conferences dedicated to ontology work in the military. The Intelligence Ontology is a resource for representing intelligence data concerning “foreign nations, hostile or potentially hostile forces or elements, or areas of actual or potential operations" and is applied to organizations, groups, and individuals. The Joint Intelligence Ontology Baseline, Military Intelligence Ontology, Intelligence Disciplines Ontology, Intelligence Organizations Ontology, Intelligence Activities Ontology, and the Intelligence Products Ontology are other examples (Figure 5.5).

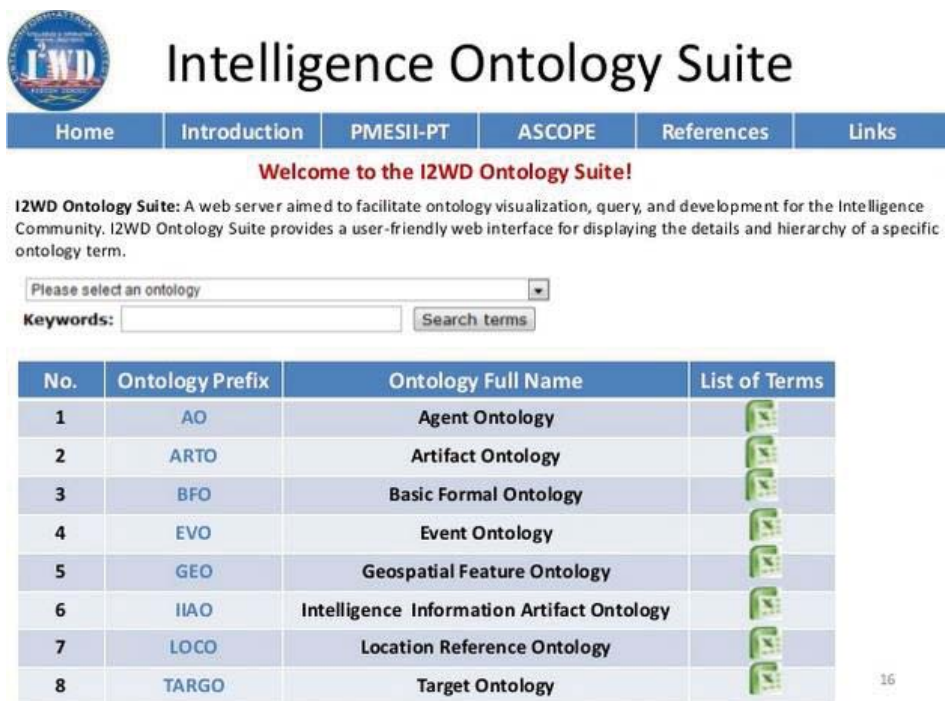


Figure 5.5 The Military's Intelligence Ontology Suite (Smith et al., 2012)

The overall aim of such ontology suits is reminiscent of the work that Ceusters produced with Manzoor (2010) in "How to Track Absolutely Everything." Systems of ontologies will allow for semantic interoperability of intelligence data and will eventually produce McConnell's vision of a globally networked and integrated intelligence enterprise. For example, the National Security Agency (NSA) is similarly involved in ontology work, as listed under the Computer & Information Sciences Research section of their website and Edward Snowden's whistleblowing. The High-Level Military Ontology, in its ability to individuate items like *Establish Civil Control Activity*, *Insurgent Cell*, *Criminal Gang*, *Social Organization*, *Mullah*, *Tribal Elder*, and *Key Leader Role* (Figures 5.6 and 5.7) are examples of the BFO being put to use in the context of social and government ontology and thus necessitate ethical review and validation.

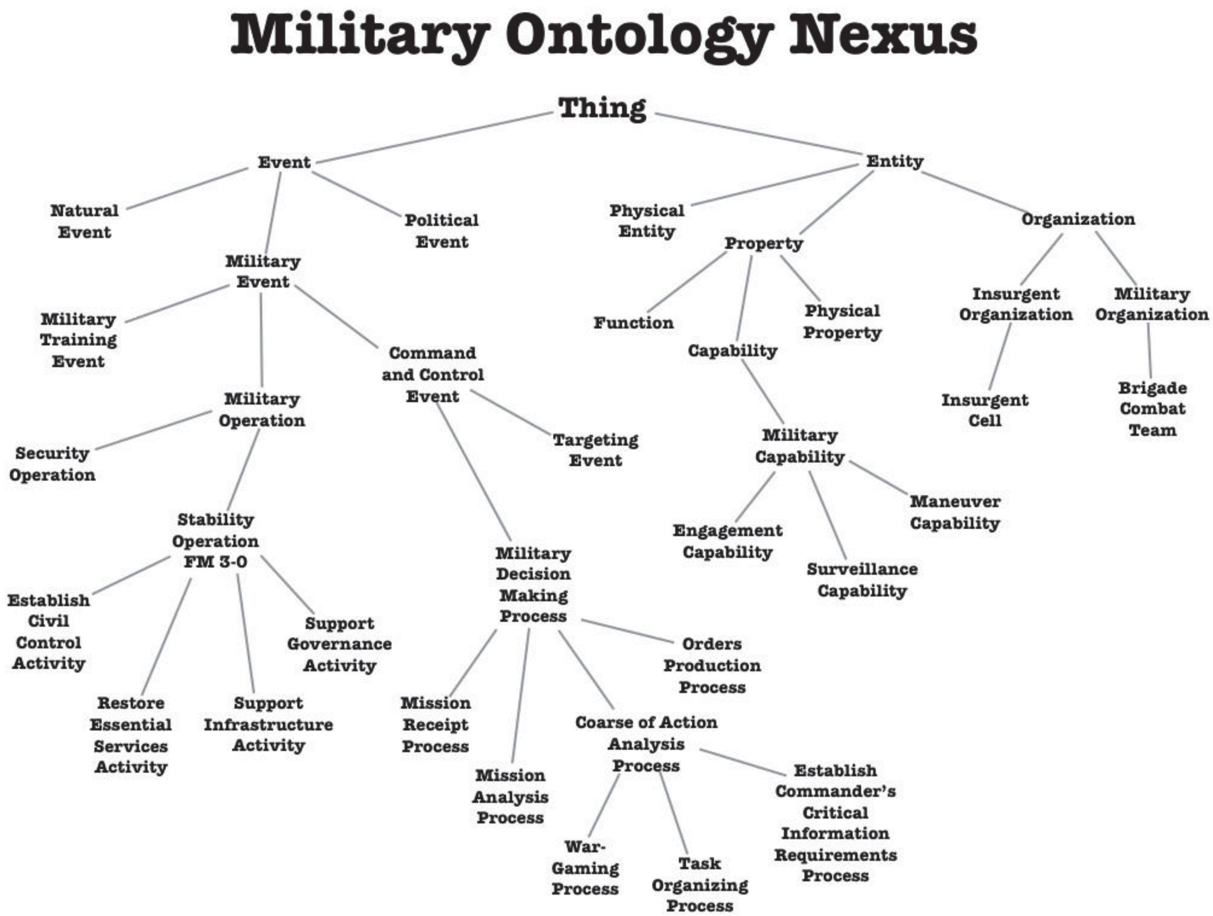


Figure 5.6 High-Level Military Ontology

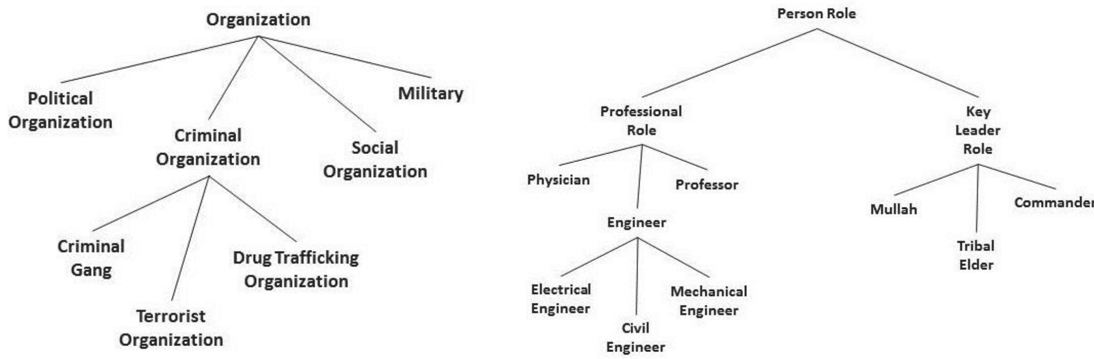


Figure 5.7 Organization and Person Role Ontologies

The types of ontologies that are being developed by the military are also being developed by agencies such as the NSA and are similarly integrating elements of social ontology into computational ontology while using the methods of ontological realism. Such practices, which concern data about things like people, places, criminal organizations, religious affiliations, political organizations, and other personal details are likely to create massive problems for citizens and the social world if such ontologies are not ethically reviewed before being applied to the social. As technologies of individuation that have the power to individuate entities and relations, those in charge of the black art of ontology building wield not only a power that can be used for the benefit of scientific progress in the domains of health and medicine but also one that can be used to potentially threaten the security and privacy of citizens.

5.5 Virtual Science

One ontologist mentioned that the end goal of building ontologies is to approach something they referred to as *virtual science*. They explained to me that the easiest way of understanding this is to think of clinical trials. At the moment, when you carry out a clinical trial, you use subjects, which may be animals, but they may also be human beings. Increasingly, it is possible to carry out clinical-trial-like processes which just use data. There are huge amounts of data already about biological phenomena, both in humans and in other organisms. It is already possible to do virtual clinical research using this data. It is already happening and part of that possibility rests on the existence of resources like the Gene Ontology, which used ontologically based ways of creating the combinability of data deriving from different kinds of sources.

There have been certain publications that have talked about data-driven science as being the end of theory. How should such provocations be read? Do we still need theories for virtual science and for data-driven science? One ontologist explained to me that they thought about this a lot. In their view, it is an empirical question. They stated that we have two different approaches to doing science or doing search on the Web, for instance. It is clear that statistically based approaches are ahead of theoretically based approaches for some kinds of problems. Google demonstrates this. Google now does use some ontology components in its work, but it is primarily still statistical and it is very successful. It works. On the other hand, when it comes to natural language understanding, the computer based on statistical approaches or computers based on semi-manual approaches to natural language understanding are still behind human beings’ capacity to understand language. To the question of science, there is a considerable amount of evidence now that computers are indispensable to many kinds of science. That does not mean that computers can do science without any kind of theory, without any kind of human input.

A computer can be programmed to extract the theoretical knowledge from a scientific paper but the computer will not do a good job. People like Smith and Ceusters are working on this problem intensely. There is a lot of funding being thrown at the problem of machine-readable science. We have not cracked this problem, even in spite of a considerable effort. We are making small steps. Some people claim that we do not need the theoretical step at all. We can do science without having something like a scientific paper with a scientific hypothesis and scientific evidence, which would then test this hypothesis and lead to improvements in the hypothesis. Some people say that we can do science in the same way that Google does search. One ontologist questioned if this could be achieved. “Whether we ever reach that point, I do not know. I think it’s an empirical question. I am sure that we’re nowhere near that point now, in spite of all the energy which has been invested in moving to that situation. We are near that point in small aspects of science, but those small aspects of science are part of science only because they are tied very closely to theory.” It is likely the work of ontological realism will continue towards achieving those aims.

One ontologist I spoke to had a different view of statistics. They told me that “In a hundred years from now we will look upon statisticians as we are now looking upon alchemists.” I wondered to what degree the same could be said of today’s emerging ontologists who are growing in numbers. Will we one day look back at the work of ontologists as one of the most important inter-scientific technologies that revolutionized the way we understand materially and informational ontology, or will they too be considered by the scientists of the future as our current alchemists engaged in the black art of ontology.

5.6 Final Remarks

This dissertation has served as a springboard into the critical study of applied computational ontology research and has provided CDS as a foundation for the future critical study of emerging technologies of individuation such as computational ontologies. The dissertation has discovered the systems of thought that inform the BFO to provide an example of how to study the theories that inform technologies of individuation and has shown how the BFO contributes to science and the engineering of social reality. The Military Ontology is only one example of applied computational ontology being put to use in domains outside of the natural sciences. In these domains, applied computational ontologies should be studied by qualitative and critical researchers and their ethical impact must be evaluated.

Preliminary Conclusions. There is a direct connection between philosophical ontology and computational ontology. The BFO's ontological realism works as shown by its widespread use. Communities of practice are important to shaping the BFO. The BFO is also applied in social, government, and other contexts.

Significance. Informational ontology has influenced scientific knowledge production and changed scientific practice and communication. Ontologies are a new way to communicate data. As technologies of individuation, applied computational ontologies like the BFO require ethical review in social ontology contexts.

Looking forward in terms of general future research, I plan on refining the theory of technologies of individuation and developing case studies based on Steinmann, Matei, and Collmann's (2016) ethical matric to measure the ethical impact of applied computational ontologies that operate in various contexts, particularly in social, government, and commerce contexts. Lastly, I plan on conducting more ethnographic field work with members working on specific iterations of the BFO such as the Military Ontology to conduct future work into applied ontology.

The communities that form around applied computational entities like the BFO observe what Slayton (2013) describes as disciplinary repertoires, which are

> the quantitative rules, codified knowledge, and habits of problem solving that enable experts to structure, estimate, and quantify uncertain technological futures. Disciplinary repertoires allow experts to rhetorically distinguish subjective, politically controversial aspects of a problem from putatively objective, technical realities (p. 2).

Such disciplinary repertories are most clearly visible in applied computational technologies like the BFO when they are put to use in individuating social reality and the individuals who make it up. The rhetorical move of labeling realism-based ontology as a purely scientific problem ignores the way in which technologies of individuation stand to impact everyday life. I plan on building on this dissertation work by examining how applied computational ontologies operate in social contexts by conducting further interviews with military ontology personnel and researchers at the Institute for Formal Ontology and Medical Information Science.

REFERENCES

REFERENCES

Allen, G. N. & March, S. T. (2012). A research note on representing part-whole relations in conceptual modeling. *MIS Quarterly, 36*(3), 945-964.

Almeida, M. B., Slaughter, L. & Brochhausen, M. (2012). Towards an ontology of document acts: Introducing a document act template for healthcare. In Herrero et al. (Eds.), *OTM 2012 Workshops* (pp. 420–425). Berlin: Springer.

Arneson, P. (2009). Ontology. In S. W. Littlejohn & K. A. Foss (Eds.), *Encyclopedia of communication theory* (pp. 695-698). Los Angeles, CA: Sage.

Arp, R., Smith, B., & Spear, A. D. (2015). *Building ontologies with basic formal ontology*. Cambridge, MA: MIT Press.

Beauchamp, T. L., & Childress, J. F. (1979/2013). *Principles of biomedical ethics*. Seventh Edition. New York, NY: Oxford University Press.

Beniger, J. (1989). *The control revolution: Technological and economic origins of the information society*. Cambridge: Harvard University Press.

Bijker, W. E., Hughes, T. P., & Pinch, T. (1987/1993). *The social construction of technological systems: New directions in the sociology and history of technology*. Cambridge: MIT Press.

Blass, A., Gurevich, Y., & Hudis, E. (2007). *The tower-of-babel problem, and security assessment sharing*. Technical Report MSR-TR-2010-57, European Association for Theoretical Computer Science.

Bodenreider, O. (2008). Biomedical ontologies in action: Role in knowledge management, data integration and decision support. *Yearbook of Medical Informatics*, 67-79.

Bodenreider, O. & Stevens, R. (2006). Bio-ontologies: Current trends and future directions. *Briefings in Bioinformatics*, *7*(3), 256-274.

Boulos, M. N. K., Yassine, A., Shirmohammadi, S., Namahoot, C. S., & Brückner, M. (2015). Towards an "internet of food": Food ontologies for the internet of things. *Future Internet*, *7*, 372-392.

Bowker, G. C. & Star, S. L. (2000). *Sorting things out: Classification and its consequences*. Cambridge, MA: MIT Press.

Bowker, G. C., Baker, K., Millerand, F., & Ribes, D. (2010). Toward information infrastructure studies: Ways of knowing in a networked environment. In J. Hunsinger, L. Klastrup, & M. Allen (Eds.), *International Handbook of Internet Research* (pp. 97-117). Dordrecht: Springer.

Boyd, D. and Crawford, K. (2012). Critical questions for big data. *Information, Communication & Society*, *15*(5), 662-79.

Bozeman, B. & Boardman, C. (2014). Research collaboration and team science: A state-of-the-art review and agenda. Cham: Springer.

Bratton, B. (2016). *The stack: On software and sovereignty*. Cambridge, MA: MIT Press.

Brodaric, B. & Gahegan, M. (2010). Ontology use for semantic e-science. *Semantic Web Journal*. 1-5.

Brown, J. & Gray, E. (1995). The people are the company. Available at: http://www.fastcompany.com/26238/people-are-company. Accessed: 19 March 2016.

Brown, J. S. & Duguid, P. (1991). Organizational learning and communities of practice: Toward a unified view of working, learning and innovation. *Organization Science*, *2*(1), 40-57.

Carmichael, D. W. (2012). *Organizing archival records: A practical method of arrangement and description for small archives*. Lanham: AltaMira Press.

Casellas, N. (2011). *Legal ontology engineering: Methodologies, modelling trends, and the ontology of professional judicial knowledge*. Dordrecht: Springer.

de Certeau, M. (1984). *The practice of everyday life*. Berkeley, CA: University of California Press.

Ceusters, W. (2006). Towards a realism-based metric for quality assurance in ontology matching. In B. Bennett & C. Fellbaum (Eds), *Formal Ontology in Information Systems. Proceedings of FOIS-2006*, Baltimore, MD (pp. 321-332), New York: IOS Press.

Ceusters W. (2012). An information artifact ontology perspective on data collections and associated representational artifacts. In J. Mantas, S. K. Andersen, M. C. Mazzoleni, B. Blobel, S. Quaglini, & A. Moen (Eds.), *Quality of life through quality of information: Proceedings of medical Informatics Europe conference* (pp. 68-72). Amsterdam: IOS Press.

Ceusters W. (2015). Ontological investigations into medical diagnoses. Grand Rounds of the UB Department of Biomedical Informatics. Wednesday, August 26, 2015, 12:00 pm – 1:00 pm, 5019A CTRC.

Ceusters, W. & Manzoor, (2010). How to track absolutely everything. In L. Obrst, T. Janssen, & W. Ceusters (Eds.), *Proceedings of the 2010 conference on ontologies and semantic technologies for intelligence* (pp. 13-36). Amsterdam: IOS Press.

Ceusters, W. & Smith B. (2005). Tracking referents in electronic health records. *Studies in Health Technology and Informatics*, 116, 71-76.

Ceusters, W. & Smith, B. (2006a). A realism-based approach to the evolution of biomedical ontologies. In *Proceedings of the Annual AMIA Symposium* (pp. 121-125). Washington, DC: American Medical Informatics Association.

Ceusters, W. & Smith, B. (2006b). Referent tracking for treatment optimisation in schizophrenic patients: A case study in applying philosophical ontology to diagnostic algorithms. *Web Semantics: Science, Services and Agents on the World Wide Web*, *4*(3), 229-236.

Ceusters, W. & Smith, B. (2007). Referent tracking for digital rights management. *International Journal of Metadata, Semantics and Ontologies*, *2*(1), 45-53.

Ceusters, W., Elkin, P., & Smith, B. (2006). Referent tracking: The problem of negative findings. *Studies in Health Technology and Informatics*, *124*, 741-6.

Chakravartty, A. (2007). *A metaphysics for scientific realism: Knowing the unobservable*. Cambridge: Cambridge University Press.

Chalmers, D. (2009). Ontological anti-realism. In D. Chalmers, D. Manley, & R. Wasserman (Eds.), *Metametaphysics: New essays on the foundations of ontology* (pp.77-129). Oxford: Oxford University Press.

Christen, K. (2009). Access and accountability: The ecology of information sharing in the digital age. *Anthropology News*, April, 4-5.

Clausewitz, C. v. (1989). *On War*. M. Howard & P. Paret (Eds.) (Trans.). Princeton: Princeton University Press.

Cocchiarella, N. (2007). *Formal ontology and conceptual realism*. Dordrecht: Springer.

Coleman, G. (2010). Ethnographic approaches to digital media. *Annual Review of Anthropology*, 39, 487-505.

Cox, A. M. (2005). What are communities of practice? A comparative review of four seminal works. *Journal of Information Science, 31*(6), 527-540.

Creswell, J. W. (2013). *Qualitative inquiry and research Design: Choosing among five approaches*. Los Angeles, CA: Sage.

Craig, R. T. (1999). Communication theory as a field. *Communication Theory*, *9*(2), 119-161.

Dafoe, A. (2015). On technological determinism: A typology, scope conditions, and a mechanism. *Science, Technology, & Human Values*, *40*(6), 1047-1076.

Dalton, C. & Thatcher, J. (2014). What does a critical data studies look like, and why do we care? Seven points for a critical approach to 'big data.' *Society & Space* open site. http://societyandspace.com/material/commentaries/craig-dalton-and-jim-thatcher-what-does-a-critical-data-studies-look-like-and-why-do-we-care-seven-points-for-a-critical-approach-to-big-data/

Davis, R., Shrobe, H., & Szolovits, P. (1993). What is a knowledge representation? *Artificial Intelligence Magazine, 14*(1), 17-33.

Depew, D. J. & Lyne, J. (2013). The productivity of scientific rhetoric. *Poroi: An interdisciplinary journal of rhetorical analysis and invention*, *9*(1), 1-20.

Dragos, V. (2013). Developing a core ontology to improve military intelligence analysis. *International Journal of Knowledge-based and Intelligent Engineering Systems, 17*, 29-36.

Eisinger, N. & Małuszynski, J. (Eds.) (2005). *Reasoning web*. Berlin: Springer.

Eschenbach, C., & Grüninger, M. (2008b). Preface. In C. Eschenbach & M. Grüninger (Eds.), *Formal ontology in information systems: Proceedings of the fifth international conference* (pp. v-vi). Amsterdam: IOS Press.

Fensel, D. (1998). *Ontologies: A silver bullet for knowledge management and electronic commerce*. Berlin: Springer.

Fetterman, D. M. (2009). *Ethnography: Step-by-step*. Los Angeles, CA: Sage.

Floridi, L. (2008a). A defence of informational structural realism. *Synthese, 161*(2), 219-252.

Floridi, L. (2008b). The method of levels of abstraction. *Minds and machines, 18*(3), 303-329.

Floridi, L. (2011). *The philosophy of information*. Oxford: Oxford University Press.

Floridi, L. (2013). *The ethics of information*. Oxford: Oxford University Press.

Floridi, L. & Illari, P. (Eds.) (2014). *The philosophy of information quality*. Cham: Springer.

Friedel, R. (2007). *A culture of improvement: Technology and the western millennium*. Cambridge, MA: MIT Press.

Gallotti, M. & Michael, J. (Eds.) (2014). *Perspectives on social ontology and social cognition*.

Geels, F. W. (2010). Ontologies, socio-technical transitions (to sustainability), and the multi-level perspective. *Research Policy*, *39*(4), 495-510.

Geertz, C. (1973). *The interpretation of cultures: Selected essays*. New York: Basic Books.

Gieryn, T. (1983). Boundary-work and the demarcation of science from non-science: Strains and interests in professional ideologies of scientists. *American Sociological Review*, *48*(6), 781-795.

Gitelman, L. (Ed.) (2013). *Raw data is an oxymoron*. Cambridge, MA: MIT Press.

Gooch, J. L. (1975). Knowing, communicating, and scientific rhetoric. *BioScience*, *25*(11), 715-717.

Grenon, P. & Smith, B. (2004). SNAP and SPAN: Towards dynamic spatial ontology. *Spatial Cognition & Computation*, *4*(1), 69-103.

Gruber, T. R. (1991). The role of common ontology in achieving sharable, reusable knowledge bases. In J. A. Allen, R. Fikes, & E. Sandewall (Eds.), *Principles of knowledge representation and reasoning: Proceedings of the second international conference* (pp. 601-602). Cambridge, MA: Morgan Kaufmann.

Gruber, T. R. (1992). *Onrolingua: A mechanism to support portable ontologies.* Technical Report KSL 91-66, Stanford University, Knowledge Systems Laboratory. Revision.

Gruber, T. R. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition*, *5*, 99-220.

Gruber, T. R. (1995). Toward principles for the design of ontologies used for knowledge sharing. *International Journal of Human-Computer Studies, 43*, 907-928.

Gruber, T. R. (2009). Ontology. In L. Liu & M. T. Özsu (Eds.), *Encyclopedia of database systems* (pp. 1963-1965). New York, NY: Springer.

Guarino, N. (1998a). Formal ontology and information systems. In N. Guarino (Ed.), Formal ontology in information systems: *Proceedings of the first international conference (FOIS'98), June 6-8, Trento, Italy*. Amsterdam: IOS Press.

Guarino, N. (Ed.) (1998b). *Proceedings of the first international conference (FOIS'98), June 6-8, Trento, Italy*. Amsterdam: IOS Press.

Guarino, N., Oberle, D., & Staab, S. (2009). What is an ontology? In S. Staab & R. Studer (Eds.), *Handbook on ontologies (2nd ed.) (pp. 1-17).* Dordrecht: Springer.

Guarino, N., & Poli, R. (1995). Editorial: The role of formal ontlogy in the information technology. *International Journal of Human-Computer Studies*, *43*(5-6), 623-624.

Guha, R.V. & Lenat, D.B. (1990). Cyc: A midterm report. *AI Magazine*, *11*(3), 32-59.

Hacking, I. (1983). *Representing and intervening: Introductory topics in the philosophy of natural science*. Cambridge: Cambridge University Press.

Haraway, D. (1991). *Simians, cyborgs and women: The reinvention of nature*. New York: Routledge.

Hayes, P. J. & McCarthy, J. (1969). Some philosophical problems from the standpoint of artificial intelligence. *Machine Intelligence*, *4*, 463-502.

Heath, T. & Bizer, C. (2011). *Linked data: Evolving the web into a global data space*.

Heiler, S. (1995). Semantic interoperability. *ACM Computing Surveys, 27*(2), 271-273.

Hepp, M. (2008). Goodrelations: An ontology for describing products and services offers onthe web. In A. Gangemi & J. Euzenat (Eds.), *Proceedings of the 16th International Conference on Knowledge Engineering and Knowledge Management* (pp. 329-346). Berlin: Springer.

Heur, B. v., Leydesdorff, L., & Wyatt, S. (2012). Turning to ontology in STS? Turning to STS through 'ontology.' *Social Studies of Science, 43*(3), 341-362.

Husserl, R. (1929/1969). *Formal and transcendental logic*. D. Cairns (Trans.). New York: Springer.

Husserl, E. (1900/2001). *Logical investigations, Vol. I*. London: Routledge.

Husserl, E. (19012001). *Logical investigations, Vol. II*. London: Routledge.

Ihde, D. (1990). *Technology and the lifeworld: From garden to earth*. Bloomington, IN: Indiana University Press.

Iliadis, A. (2013). Informational ontology: The meaning of Gilbert Simondon's concept of individuation. *communication +1*, *2*(1), 1-19.

Iliadis, A. & Russo, F. (forthcoming). Critical data studies. *Big Data & Society*.

Ingarden, R. (1964). *Time and modes of being*. H. R. Michejda (Trans.). Springfield, Illinois: Charles C. Thomas.

Ingarden, R. (1947/2013). *Controversy over the existence of the world*. A. Szylewicz (Trans.). Frankfurt: Peter Lang.

Jakus, G., Milutinović, V., Omerović, S. & Tomažič, S. (2013). *Concepts, ontologies, and knowledge representation*. New York, NY: Springer.

Kallinikos, J., Aaltonen, A., & Marton, A. (2013). The ambivalent ontology of digital artifacts. *MIS Quarterly*, *37*(2), 357-370.

Kaltenbrunner, W. (2015). Infrastructural inversion as a generative resource in digital scholarship. *Science as Culture*, *24*(1), 1-23.

Katz, J. J. (1966). *The philosophy of language*. New York, NY: Harper and Row.

Katz, J. J. (1981). *Language and other abstract objects*. Totowa, NJ: Rowman and Littlefield.

Kifer, M. (2005). Rules and ontologies in f-logic. In N. Eisinger & J. Małuszynski (Eds.), *Reasoning Web* (pp. 22-34). Berlin: Springer.

Kitchin, R. (2014). *The data revolution: Big data, open data, data infrastructures and their consequences*. Los Angeles, CA: Sage.

Kitchin, R. & Lauriault, T. P. (2014). Towards critical data studies: Charting and unpacking data assemblages and their work. The Programmable City Working Paper 2.

Kitchin, R. & McArdle, G. (2016). What makes big data, big data? Exploring the ontological characteristics of 26 datasets. *Big Data & Society*, 1-10.

Latour, B. (1988). *The pasteurization of France*. A Sheridan & J Law (Trans.). Cambridge, MA: Harvard University Press.

Latour, B. & Woolgar, S. (1979). *Laboratory life: The construction of scientific facts*. Princeton, NJ: Princeton University Press.

Lave, J. & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge: Cambridge University Press.

Lenat, D.B. & Guha, R.V. (1990). *Building large knowledge based systems*. Reading, MA: Addison Wesley.

Mandrick, W. S. (2008). The ontology of counterinsurgency. Annual meeting of the MPSA Annual National Conference, Palmer House Hotel, Hilton, Chicago, IL.

Manzoor, S., Ceusters, W., & Smith, B. (2009). Referent tracking for command and control messaging systems. Ontology for the Intelligence Community, Fairfax Virginia, October 21-22, 2009.

Marquardt, W., Morbach, J., Wiesner, A. & Yang, A. (2010). *Ontocape: A re-usable ontology for chemical process engineering*. Berlin: Springer.

Mayor, C. & Robinson, L. (2014). Ontological realism and classification: Structures and concepts in the gene ontology. *Journal of the Association for Information Science and Technology*, *66*(4), 686 – 697.

McCarthy, J. (1980). Circumscription—A form of non-monotonic reasoning. *Artificial Intelligence*, *13*(1-2), 27-39.

McPhee, R. D. & Zaug, P. (2000). The communicative constitution of organizations: A framework for explanation. *Electronic Journal of Communication/La Revue Electronique de Communication*, *10*(1-2).

Merrill, G. H. (2010). Ontological realism: Methodology or misdirection? *Applied Ontology*, *5*, 79-108.

Merrill, G. H. (2011). Ontology, ontologies, and science. *Topoi*, *30*, 71-83.

Michel, A. (2016). The mutual constitution of persons and organizations: An ontological perspective on organizational change. *Organization Science*, *25*(4), 1082-1110.

Mouton, F., Leenen, L., Malan, M. M., & Venter, H. S. (2014). Towards an ontological model defining the social engineering domain. In K. Kimppa, D. Whitehouse, T. Kuusela, & J. Phahlamohlaka (Eds.), *ICT and eociety 11th IFIP TC 9 International Conference on human choice and computers, HCC11 2014 Turku, Finland, July 30 – August 1, 2014 proceedings* (pp. 266-279). Heidelberg: Springer.

Munn, K., & Smith, B. (Eds.) (2008). *Applied ontology: An introduction*. Frankfurt: Ontos Verlag.

Niles, I. & Pease, A. (2001). Towards a standard upper ontology. In C. Welty and B. Smith (Eds.), *Proceedings of the 2nd international conference on formal ontology in information systems* (pp. 2-9). ACM Press.

Nissenbaum, H. (2011). *Privacy in context technology, policy, and the integrity of social life*. Stanford, CA: Stanford University Press.

Noy, N. F., & McGuinness, D. L. (2001). *Ontology development 101: A guide to creating your first ontology*. Stanford knowledge systems laboratory technical report KSL-01-05 and Stanford medical informatics technical report SMI-2001-0880.

Oderberg, D. S. (Ed.) (2013). *Classifying reality*. Malden, MA: Wiley-Blackwell.

Office of the Director of National Intelligence. (2008). Vision 2015: A Globally Networked and Integrated Intelligence Enterprise. http://www.au.af.mil/au/awc/awcgate/dni/vision_2015_july08.pdf

Okada, M. & Smith, B. (Eds.) (2008). *Proceedings of the first interdisciplinary ontology meeting (InterOntology 2008)*. Tokyo: Keio University Press.

Orozco, J. M. S. (2012). *Applied ontology engineering in cloud services, networks and management systems*. New York: Springer.

Pesquita, C. & Couto, F. M. (2012). Predicting the extension of biomedical ontologies. *PLOS Computational Biology*, *8*(9), 1-16.

Pileggi, S. F. & Fernandez-Llatas, C. (2012). *Semantic interoperability: Issues, solutions, challenges*. Aalborg: River Publishers.

Poli, R. (1993). Husserl's conception of formal ontology. *History and Philosophy of Logic*, *14*(1), 1-14.

Poli, R. (1996). Ontology for knowledge organization. In R. Green (Ed.), *Knowledge organization and change: Proceedings of the Fourth International ISKO Conference* (pp. 313-319). Frankfurt: Indeks.

Poli, R. (2001). The basic problem of the theory of levels of reality. *Axiomathes*, *12*, 261-283.

Poli, R., Healy, M., & Kameas, A. (Eds.) (2010). *Theory and applications of ontology: computer applications*. Dordrecht: Springer.

Putnam, L. L. & Nicotera, A. M. (Eds.) (2009). *Building theories of organization: The constitutive role of communication*. New York, NY: Routledge.

Putnam, L. L. & Nicotera, A. M. (2010). Communicative constitution of organization is a question: Critical issues for addressing it. *Management Communication Quarterly*, *24*(1), 158-165.

Putnam, L. L., Nicotera, A. M. & McPhee, R. D. (2009). Introduction: Communication constitutes organization. In L. L. Putnam & A. M. Nicotera (Eds.), *Building theories of organization: The constitutive role of communication* (pp. 1-19). New York, NY: Routledge.

Quine, W.V. (1948). On what there is. *Review of Metaphysics*, *2*, 21-38.

Quine, W.V. (1960). *Word and object*. Cambridge, MA: MIT Press.

Ramsey, A. E., Sharer, W. B., L'Eplattenier, B., Mastrangelo, L. S. (Eds.) (2009). *Working in the archives practical research methods for rhetoric and compositio*n. Carbondale, IL: Southern Illinois University Press.

Raskin, V., Taylor, J. M., & Hempelmann, C. F. (2010). Ontological semantic technology for detecting insider threat and social engineering. In R. Ford & C. Gates (Eds.), *Proceedings of the 2010 workshop on new security paradigms* (pp. 115-128). New York, NY: ACM.

Rector, A. L., Nolan, W., A., & Kay, S. (1991). Foundations for an electronic medical record. *Methods of Information in Medicine*, 30, 179-86.

Rheingold, H. (1993). *The virtual community: Homesteading on the electronic frontier*. New York, NY: Harper Perennial.

Riaño, D. (Ed.) (2009). *Knowledge management for health care procedures*. Berlin: Springer.

Rector, A.L., Nolan, W.A., & Kay, S. (1991). Foundations for an electronic medical record. *Methods of Information in Medicine, 30*, 179-86.

Rogers, R. (2013). *Digital methods*. Cambridge, MA: MIT Press.

Roszak, T. (1986/1994). *The cult of information: A neo-luddite treatise on high-tech, artificial intelligence, and the true art of thinking*. Oakland: University of California Press.

Roussey, C., Pinet, F., Kang, M. A., & Corcho, O. (2011). An Introduction to Ontologies and Ontology Engineering. In G. Falquet, C. Métral, J. Teller, & C. Tweed (Eds.), *Ontologies in urban development projects* (pp. 9-38). Berlin: Springer.

Santos, G. C. (2015). Ontological emergence: How is that possible? Towards a new relational ontology. *Foundations of Science*, *20*(4), 429-446.

Schiavone, F. (2014). *Communities of practice and Vintage innovation: A strategic reaction to technological change*. Cham: Springer.

Schober, D., Smith, B., Lewis, S. E., Kusnierczyk, W., Lomax, J., Mungall, C., Taylor, C. F., Rocca-Serra, P., & Sansone, S-A. (2009). Survey-based naming conventions for use in OBO Foundry ontology development. *BMC Bioinformatics*, *10*(125).

Schulze-Kremer, S. (2002). Ontologies for molecular biology and bioinformatics. *In Silico Biology*, *2*(3), 179-193.

Searle, J (1995). *The construction of social reality*. New York, NY: The Free Press.

Searle, J. (2006). Social ontology: Some basic principles. *Anthropological Theory*, *6*(1), 12-29.

Simondon, G. (2005). *L'Individuation à la lumière des notions de forme et d'information*. Paris: Jérôme Millon.

Sivamani, S., Bae, N., & Cho, Y. (2013). A smart service model based on ubiquitous sensor networks using vertical farm ontology. *International Journal of Distributed Sensor Networks*, *2013*, 1-8.

Slayton, R. (2013). *Arguments that count: Physics, computing, and missile defense, 1949–2012*. Cambridge, MA: MIT Press.

Smith, B. (1995a). Towards an ontology of common sense. In J. Hintikka & K. Puhl (Eds.), *The British tradition in 20th century philosophy: Proceedings of the 17th international Wittgenstein symposium* (pp. 300-308). Vienna: Hölder-Pichler-Tempsky.

Smith, B. (1995b). Formal ontology, common sense and cognitive science. *International Journal of Human-Computer Studies*, *43*, 641-667.

Smith, B. (2001). Fiat objects. *Topoi*, *20*(2), 131-148.

Smith, B. (2003). Towards industrial-strength philosophy: How analytical ontology can help medical informatics. *Interdisciplinary Science Reviews*, *28*, 106-111.

Smith, B. (2004a). Ontology. In L. Floridi (Ed.), *The blackwell guide to the philosophy of computing and information* (pp. 155-166). Oxford: Blackwell.

Smith, B. (2004b). Beyond concepts: ontology as reality representation. In A. Varzi & L. Vieu (Eds.), *International conference on formal ontology and information: Proceedings of the third international conference* (pp. 73-84). Turin: IOS Press.

Smith, B. (2005). Against fantology. In M.E. Reicher & J.C. Marek (Eds), *Experience and analysis: Papers of the 27th International Wittgenstein Symposium*. Vienna: The Austrian Ludwig Wittgenstein Society.

Smith, B. (2006). From concepts to clinical reality: An essay on the benchmarking of biomedical terminologies. *Journal of Biomedical Informatics*, *39*, 288-298.

Smith, B. (2009). From Aristotle to the Universal Core. Conference Presentation. Buffalo.

Smith, B. (2014). Document acts. In A. Konzelmann-Ziv & H. B. Schmid (Eds.), *Institutions, emotions, and group agents: Contributions to social ontology* (pp. 19-31). Dordrecht: Springer.

Smith, B. (2016). Towards a science of emerging media. In J. E. Katz and J. Floyd (Eds.), *Philosophy of emerging media: Understanding, appreciation and application* (pp. 29-48). Oxford: Oxford University Press.

Smith, B. & Ceusters W. (2010). Ontological realism: A methodology for coordinated evolution of scientific ontologies. *Applied Ontology*, *5*(3-4), 139-188.

Smith, B. & Smith, D. W. (Eds.) (1995a). *The Cambridge companion to Husserl*. Cambridge: Cambridge University Press.

Smith, B. & Smith, D. W. (1995b). Introduction. In S. Smith & D. W. Smith (Eds.), *The Cambridge companion to Husserl* (pp. 1-44). Cambridge: Cambridge University Press.

Smith, B., Malyuta, T., Mandrick, W. S., Fu, C., Parent, K., & Patel, M. (2012). Horizontal integration of warfighter intelligence data: A shared semantic resource for the intelligence community. In P. C. G. Costa & K. B. Laskey (Eds.), *Proceedings of the seventh international conference on semantic technologies for intelligence, defense, and security* (pp. 112-119). CEUR Workshop Proceedings.

Smith, B. & Welty, C, (2001). Ontology: Towards a new synthesis. In B. Smith & C. Welty (Eds.), *Proceedings of the international conference on formal ontology in information systems* (pp. 3-9). New York, NY: ACM.

Smith, B., Ashburner, M., Rosse, C., Bard, J., Bug, W., Ceusters, W., Goldberg, L. J., Eilbeck, K., Ireland, A., Mungall, C. J., the OBI Consortium, Leontis, N., Rocca-Serra, P., Ruttenberg, A., Sansone, S-A, Scheuermann, R. H., Shah, N., Whetzel, P. L., & Lewis, S. (2007). The OBO Foundry: Coordinated evolution of ontologies to support biomedical data integration. *Nature Biotechnology*, *25*, 1251-1255.

Smith, B. C. (1996). *On the origin of objects*. Cambridge, MA: MIT Press.

Staab, S. & Studer, R. (Eds.) (2009). *Handbook on ontologies (2nd ed.)*. Dordrecht: Springer.

Star, S. L. (1999). The ethnography of infrastructure. *American Behavioral Scientist*, *43*(3), 377-391.

Star, S. L. & Lampland, M. (Eds.) (2009). *Standards and their stories: How quantifying, classifying, and formalizing practices shape everyday life*. Ithaca, NY: Cornell University Press.

Steinmann, M., Matei, S. A., & Collmann, J. (2016). A theoretical framework for ethical reflection in big data research. In J. Collmann & S. A. Matei (Eds.), *Ethical reasoning in big data*. Cham: Springer.

Stevens, H. (2013). *Life out of sequence: A data-driven history of bioinformatics*. Chicago, IL: University of Chicago Press.

Suárez-Figueroa, M. C., Gómez-Pérez, A., Motta, E. & Gangemi, A. (2012a). *Ontology engineering in a networked world*. Heidelberg: Springer.

Suárez-Figueroa, M. C., Gómez-Pérez, A., Motta, E. & Gangemi, A. (2012b). Introduction: Ontology engineering in a networked world. In M. C. Suárez-Figueroa, A. Gómez-Pérez, E. Motta, & A. Gangemi (Eds.), *Ontology engineering in a networked world* (pp. 1-6). Heidelberg: Springer.

Uschold, M., & Gruninger, M. (1996). Ontologies: Principles, methods and applications. *The Knowledge Engineering Review*, *11*(2), 93-136.

Wagner-Pacifici, R., Mohr, J. W., & Breiger, R. L. (2015). Ontologies, methodologies, and new uses of Big Data in the social and cultural sciences. *Big Data & Society*, 1-11.

Warren, P., Mulholland, P., Collins, T., & Motta, E. (2014). Using ontologies understanding the user experience. *Lecture Notes in Computer Science*, *8876,* 579-590.

Weisberg, M. (2013). *Simulation and similarity: Using models to understand the world*. Oxford: Oxford University Press.

Wenger, E. (1998). *Communities of practice: Learning, meaning, and identity*. Cambridge: Cambridge University Press.

Wenger, E., McDermott, R. & Snyder, W. M. (2002). *Cultivating communities of practice*. Boston, MA: Harvard Business School Press.

White, L. (1966). *Medieval technology and social change*. Oxford: Oxford University Press.

Winner, L. (1986/1988). *The whale and the reactor: A search for limits in an age of high technology*. Chicago: University Of Chicago Press.

Weiss, R. S. (1995). *Learning from strangers: The art and method of qualitative interview studies*. New York: NY. Simon and Schuster.

Williams, R., & Edge, D. (1996). The social shaping of technology. *Research Policy*, *25*, 865-899.

Woolgar, S. & Lezaun, J. (2013). The wrong bin bag: A turn to ontology in science and technology studies? *Social Studies of Science*, *43*(3), 321-340.

Yim, P. P. (2015). Bootstrapping the applied ontology practice: Ontology communities, then and now. *Applied Ontology*, *10*(3-4), 229-241.

Zemach, E. (1970). Four ontologies. *The Journal of Philosophy*, *67*(8), 231-247.

Ziv, A. K. & Schmid, H. B. (Eds.) (2014). *Institutions, emotions, and group agents: Contributions to social ontology*. Dordrecht: Springer.

Zúñiga, G. L. (2001). Ontology: Its transformation from philosophy to information systems. In N. Guarino, B. Smith, & C. Welty (Eds.), *Proceedings of the international conference on formal ontology in information systems* (pp. 187-197). New York, NY: IOS Press.

APPENDICES

Appendix A Interview Questions

1. What is the difference between an ontologist and a philosopher who studies ontology?

2. What is ontology (in philosophy and computers)?

3. Can you think of any philosophers in history who engaged in ontology?

4. Historically, which philosophers or books would you consider as important to ontology?

5. Is computational ontology related to philosophical ontology?

6. What are some of the unique problems in big data that philosophy can help solve?

7. Why is ontology important to biomedical informatics?

8. What other fields is ontology important to?

9. What is virtual science?

10. Are we entering a stage of data driven science and the end of theory?

11. What is the New York State Center of Excellence in Bioinformatics & Life Sciences?

12. Can you tell me about the National Center for Ontological Research?

13. Can you tell us a little about contemporary ontology work?

14. What is Basic Formal Ontology?

15. Can Basic Formal Ontology be applied in any domain?

16. Is ontology a matter of organizational communication? Is it philosophy?

17. What is the Open Biomedical Ontologies consortium?

18. What is the Ontology for Biomedical Investigations?

19. What is the Information Artifact Ontology?

20. Can you tell me about Building Ontologies with Basic Formal Ontology?

21. How do computer scientists work with philosophers?

22. How the actual ontologies constructed and who builds them?

23. How are ontology research teams set up?

24. What do you view as the biggest challenge to ontology today?

25. What is the future of ontological research?

26. Have you encountered any resistance or critics of ontology?

27. What areas are underdeveloped in ontology?

28. Where can I find examples of these ontologies?

Appendix B BFO Ontologies and their Contexts

Ontology	Social	Government	Commerce	Science
ACGT Master Ontology				x
Adverse Event Ontology				x
Adverse Event Reporting Ontology				x
AFO Foundational Ontology	x	x	x	x
Actionable Intelligence Retrieval System	x	x		
Bacterial Clinical Infectious Diseases Ontology				x
Behavior Perspective Model	x		x	
Bank Ontology			x	
Beta Cell Genomics Application Ontology				x
BioAssay Ontology				x
Bioinformatics Web Service Ontology				x
Biological Collections Ontology				x
Biomedical Ethics Ontology				x
Biomedical Grid Terminology				x
BioTop: A Biomedical Top-Domain Ontology				x
BIRNLex				x
Blood Ontology				x
Body Fluids Ontology				x
Bone Dysplasia Ontology				x
Cancer Cell Ontology				x
Cancer Chemoprevention Ontology				x
Cardiovascular Disease Ontology				x
Cell Behavior Ontology				x
Cell Cycle Ontology				x
Cell Expression, Localization, Development and Anatomy Ontology				x

Cell Line Ontology				x
Cell Ontology				x
Chemical Analysis Ontology				x
Chemical Entities of Biological Interest				x
CHRONIOUS Ontology Suite				x
Cigarette Smoke Exposure Ontology				x
Clusters of Orthologous Groups (COG) Analysis Ontology				x
Cognitive Paradigm Ontology	x			x
Common Anatomy Reference Ontology				x
Communication Standards Ontology	x	x	x	x
Conceptual Model Ontology	x	x	x	x
Coriell Cell Line Ontology				x
CPR Ontology				x
Document Act Ontology	x	x	x	x
Drug Interaction Ontology				x
Drug Ontology				x
Drug-drug Interaction Evidence Ontology				x
Drug-drug Interaction Ontology				x
Dynamic Earth Sciences Ontologies		x		x
Eagle-I Research Resource Ontology				x
Economics Ontology	x		x	
Email Ontology	x			
Emotion Ontology	x			
Environment Ontology		x		x
Epidemiology Ontology	x			x
Epilepsy and Seizure Ontology				x
Evolution Ontology				x
Experimental Factor Ontology				x

(EXperimental ACTioins) Biomedical Protocoll Ontology				x
Exposé: An Ontology for Data Mining Experiments	x	x	x	x
Flybase Drosophila Anatomy Ontology				x
Fission Yeast Phenotype Ontology				x
Flower-Visiting Domain Ontology				x
Ontology and Observation-Date Application Ontology				x
Foundational Model of Anatomy				x
Gastrointestinal Endoscopy Ontology				x
Gene Regulation Ontology				x
General Information Model	x	x	x	x
Genomic Feature and Variation Ontology				x
Gestalt				x
Health Data Ontology Trunk				x
Human Interaction Network Ontology				x
Human Physiology Simulation Ontology				x
Infectious Disease Ontology				x
Information Artifact Ontology	x	x	x	x
Informed Consent Ontology				x
Interaction Network Ontology				x
Interdisciplinary Prostate Ontology Project				x
Intracranial aneurysm Ontology				x
Knowledge Base Of Biomedicine				x
Lipid Ontology				x
Materials Ontology				x
Mental Disease Ontology				x
Mental Functioning Ontology				x

miRNAO				x
Middle Layer Ontology for Clinical Care				x
Military Ontology	x	x		
MIRO and IRbase				x
Model for Clinical Information				x
Mouse Pathology Ontology				x
Nanoparticle Ontology				x
NeuroPsychological Testing Ontology				x
Neuroscience Information Framework				x
Neuroscience Information Framework Standard Ontology				x
Neural Electromagnetic Ontologies				x
Neuroscience Information Framework				x
Neuroscience Information Framework Subcellular Ontology				x
New Upper Level Ontology	x	x	x	x
Non-Coding RNA Ontology				x
NMR-Instrument Component of Metabolomics Investigations Ontology				x
Ocular Disease Ontology				x
OncoCL-KB				x
OntoAlign++	x	x	x	x
OntoForInfoScience				x
Ontologized Minimum Information About BIobank data Sharing				x
Ontology for Autism Spectrum Disorder				x
Ontology for Biomedical Investigations				x
Ontology for Dengue Fever				x
Ontology for Drug Discovery Investigations			x	x

Ontology for Energy Investigations		x	x	x
Ontology for General Medical Science				x
Ontology for Genes and Genomes				x
Ontology for Genetic Interval				x
Ontology for Genetic Susceptibility Factor				x
Ontology for Guiding Appropriate Antibiotic Prescribing				x
Ontology for Laparoscopic Surgeries				x
Ontology for Microbial Phenotypes				x
Ontology for MIcroRNA Target Prediction				x
Ontology for Newborn Screening and Translational Research				x
Ontology for Next Generation Sequencing Experiments				x
Ontology for Pain and Related Disability, Mental Health and Quality of Life	x			x
Ontology for Periodontitis				x
Ontology of Clinical Research				x
Ontology of Biobanking Administration				x
Ontology for Parasite LifeCycle				x
Ontology for Rehabilitation				x
Ontology of Biological and Clinical Statistics				x
Ontology of Data Mining			x	x
Ontology of Medically Related Social Entities	x			x
Ontology of Social Participation	x			
Ontology of Vaccine Adverse Events				x
Ontology-Based Data Access			x	x
Ontology-Based eXtensible Data Model				x

Ontology-Driven Information System/Ontology-Driven Scenario Generator				x
Oral Health and Disease Ontology				x
Parasite Experiment Ontology				x
Patient Safetry Categorial Structure				x
Petrochemical Ontology				x
Phenotypic Quality Ontology				x
Plant Ontology				x
Population and Community Ontology	x			
Proper Name Ontology	x	x	x	x
Protein-Ligand Interaction Ontology				x
Proteomics data and process provenance ontology				x
Protein Ontology				x
RNA Ontology				x
Schistosomiasis Process Ontology				x
Saliva Ontology				x
Semantic EHR Model from Linked2Safety				x
Senselab Ontology				x
Sequence Ontology				x
Semanticscience Integrated Ontology				x
Sleep Domain Ontology				x
SMART Protocols	x	x	x	x
Spatiotemporal Ontology for the Administrative Units of Switzerland		x		
Special Nuclear Materials Detection Ontology				x
Statistics Ontology	x	x	x	x
Subcellular Anatomy Ontology of NCMIR				x
Suggested Ontology for Pharmacogenomics				x
Taxonomy for rehabilitation of knee conditions				x

Time Event Ontology	x	x	x	x
Translational Medicine Ontology				x
(Microbial) Typing Ontology				x
United Nations SDG Ontology Framework		x		
Universal Core Semantic Layer	x	x	x	x
Vaccine Ontology				x
Xenopus Anatomy Ontology				x
YAMATO	x	x	x	x
yOWL				x
Zebrafish Anatomical Ontology				x

VITA

VITA

ANDREW ILIADIS

Purdue University
Brian Lamb School of Communication
Beering Hall of Liberal Arts and Education, Room 2114
100 N. University Street, West Lafayette, IN 47907-2098
Main Office: (765) 494-3429 / Mobile: (765) 714-2651
FAX: (765) 496-1394 / ailiadis@purdue.edu
www.andrewiliadis.com

Academic Appointments

Instructor, Purdue University 2013-present
Brian Lamb School of Communication
Instructor, Purdue University 2015-present
Department of Philosophy
Instructor, Purdue University 2015-present
Department of Computer Science

Education

Ph.D., Purdue University 2012-2016 (expected)
Communication and Philosophy, Lafayette, IN, USA
Title: *A Black Art: Ontology, Data, and the Tower of Babel Problem*
Co-Advisors: Ashley R. Kelly and Daniel W. Smith
Committee: Sorin A. Matei and Daniel R. Kelly

M.A., Ryerson University and York University 2008-2010
Communication and Culture, Toronto, ON, Canada
Advisor: Stuart J. Murray
Committee: Alan D. Sears and Colin Mooers

B.A., Trent University 2003-2007
English Literature and Cultural Studies, Peterborough, ON, Canada
Dean's Honor Roll

Summer School Fellowship, Université de Paris (Ouest Nanterre) 2014
Seminar Leaders: Jean-Michel Salanskis and Élie During

Edited Collections

Iliadis, A. (Ed.) (2015). Book symposium on *Le concept d'information dans la science contemporaine*. *Philosophy & Technology*. Springer. Online first (June 30). http://goo.gl/0LS4q0

Iliadis, A. and Russo, F. (Eds.) (forthcoming). Critical Data Studies. *Big Data & Society*. Sage. http://goo.gl/wLBoSN

Iliadis, A. (Ed.) (forthcoming). Gilbert Simondon and Raymond Ruyer. *Deleuze Studies*. Edinburgh University Press. https://goo.gl/Zr4hw6

Peer-Reviewed Publications

Journal Articles

Iliadis, A. (2015). The right to nonparticipation for global digital citizenship. *International Review of Information Ethics*, *23*, 20-34. http://goo.gl/FBDE0n

Iliadis, A. (2015). Mechanology: Machine typologies and the birth of philosophy of technology in France (1932-1958). *Systema*, *3*(1), 119-130. http://goo.gl/E4jEeq

Barthélémy, J-H and **Iliadis, A.** (2015). Gilbert Simondon and the philosophy of information: An interview with Jean-Hugues Barthélémy. *Journal of French and Francophone Philosophy*, *23*(1), 102-112. http://goo.gl/OHZbxh

Iliadis, A. (2015). The concept of information in contemporary science (Royaumont, 1962). *Philosophy & Technology*, online first (June 30), 1-3. http://goo.gl/nEXGwh

Iliadis, A. (2015). Two examples of concretization. *Platform: Journal of Media and Communication*, *6*(1), 86-95. https://goo.gl/V0Bmb5
*Spanish. Bernabé Ferreyra (Trans.) https://goo.gl/RbFDCo

Iliadis, A. (2013). A new individuation: Deleuze's Simondon connection. *MediaTropes*, *4*(1), 83-100. http://goo.gl/rHExEf

Iliadis, A. (2013). Informational ontology: The meaning of Gilbert Simondon's concept of individuation. *communication +1*, *2*(1), 1-19. http://goo.gl/PmwkAw

Iliadis, A. (2010). The 'claws of absolute necessity': Deleuze on culture. *eTopia*, on-line initiative of *TOPIA: Canadian Journal of Cultural Studies*. http://goo.gl/nwCRAJ

Proceedings and Reports

Brightman, A., Beaver, J., Hess, J., **Iliadis, A.**, Kisselburgh, L., Krane, M., Loui, M., Zoltowski, C. (2016). PRIME ethics: Purdue's reflective & interactive modules for engineering ethics. *Infusing ethics into the development of engineers: Exemplary education activities and programs*. Infusing Ethics Selection Committee, Center for Engineering Ethics and Society, National Academy of Engineering. Washington, DC: The National Academic Press. http://goo.gl/rmMOKi

Hess, J., Beever, J., **Iliadis, A.**, Kisselburgh, L., Zoltowski, C., Krane, M., Brightman, A. (2014). An ethics transfer case assessment tool for measuring ethical reasoning abilities of engineering students using reflexive principlism approach. *Frontiers in education conference proceedings* (pp. 1-5). Madrid: IEEE. http://goo.gl/fbYxCa

Kisselburgh, L., Zoltowski, C., Beever, J., Hess, J., **Iliadis, A.**, Krane, M., Brightman, A. (2014). Effectively engaging engineers in ethical reasoning about emerging technologies: A cyber-enabled framework of scaffolded, integrated, and reflexive analysis of cases. *American society for engineering education conference proceedings* (pp. 1-17). Indianapolis: ASEE. http://goo.gl/YuXDk6

Book Chapters

Iliadis, A. (2015). A quick history of the philosophy of information. In Phyllis Illari (Ed.), *The philosophy of information: A simple introduction (pp. 9-27)*. London: Society for the Philosophy of Information. http://goo.gl/eV1iR7

Translations

Simondon, N. (2015). Gilbert Simondon and the Royaumont colloquium, 1962. *Philosophy & Technology*, online first (June 30), 16-23. http://goo.gl/W579Ew

Garcia, T. (2014). Interview with Tristan Garcia. *Figure/Ground*. September 28th, 2014. http://goo.gl/kuQGpj

Book Reviews

Iliadis, A. (2015). Media effects research: A basic overview. *Canadian Journal of Communication*, *40*(3), 582-584. http://goo.gl/SPMq5p

Iliadis, A. (2015). *Why philosophize?* by Jean-François Lyotard. *Dialogue: Canadian Philosophical Review*, online first (May 25), 1-2. http://goo.gl/Q8EllD

Iliadis, A. (2015). Big data. *Communication Booknotes Quarterly*, *46*(2), 54-57. http://goo.gl/xyPMx8

Iliadis, A. (2013). *The uprising: On poetry and finance* by Franco 'Bifo' Berardi. *Marx & Philosophy Review of Books*, online first (July 30). http://goo.gl/dDxY4A

Journalism and Literary Publications

Iliadis, A. (2015). IDRS Review – Session 4 with Kevin Brock, Ashley R. Kelly, Annette Vee, and James J. Brown Jr. Digital Rhetoric Collaborative. http://goo.gl/H9Uz29

Iliadis, A. (2014). Interview with Iris van der Tuin. *Figure/Ground*. http://goo.gl/uqtvlp

Iliadis, A. (2014). The informational sublime. *Ethical Technology*. Institute for Ethics and Emerging Technologies. http://goo.gl/c9n3ma

Iliadis, A. (2013). Embodied cognition's philosophical roots. *Ethical Technology*. Institute for Ethics and Emerging Technologies. http://goo.gl/9PjgCn

Iliadis, A. (2013). Interview with Vincent Mosco. *Figure/Ground*. http://goo.gl/pLnj2W

Iliadis, A. (2013). Interview with Graham Harman. *Figure/Ground*. http://goo.gl/22f8R6

Iliadis, A. (2013). Interview with Bruno Latour. *Figure/Ground*. http://goo.gl/GB1Z3j
*Spanish: *Razón y Palabra*, *18*(84), 1-8. http://goo.gl/5n96tD

Iliadis, A. (2013). Interview with Joanna Zylinska. *Figure/Ground*. http://goo.gl/Mn4EN1

Iliadis, A. (2013). Interview with Barry Wellman. *Figure/Ground*. http://goo.gl/2KLGGM

Iliadis, A. (2013). Interview with Nina Power. *Figure/Ground*. http://goo.gl/iWPVGh

Iliadis, A. (2013). Interview with Peter Adamson. *Figure/Ground*. http://goo.gl/gCVyOG

Iliadis, A. (2013). Interview with Steven Shaviro. *Figure/Ground*. http://goo.gl/l5hx5S

Iliadis, A. (2013). Interview with Gary Genosko. *Figure/Ground*. http://goo.gl/OD5znc

Iliadis, A. (2013). Interview with John Durham Peters. *Figure/Ground*. http://goo.gl/EJ2UFx

Iliadis, A. (2012). Interview with McKenzie Wark. *Figure/Ground*. http://goo.gl/KVaKyY

Iliadis, A. (2012). Interview with John Searle. *Figure/Ground*. http://goo.gl/2P84EB
*Spanish: *Razón y Palabra*, 83, Junio-Agosto. http://goo.gl/SNxkiJ

Iliadis, A. (2012). "Interview with Jodie Dean. *Figure/Ground*. http://goo.gl/D5LAZR

Iliadis, A. (2012). "Interview with Mark Kingwell. *Figure/Ground*. http://goo.gl/HCUHo0

Iliadis, A. (2010). “i-everything: Mapping Ryerson’s Neoliberal University.” *York University Free Press*.

Iliadis, A. (2003). Shepherd of sharks. *3:AM Magazine*. http://goo.gl/VzuvW1

Awards

Social Sciences and Humanities Research Council of Canada Postdoctoral Fellowship (waitlist)

National Academy of Engineering Exemplar in Engineering Ethics Education, 2016.

Decimal Lab Doctoral Fellowship, $5,320, University of Ontario Institute of Technology, 2016.

5th Annual International Symposium on Digital Ethics, Top Student Paper Award, Loyola University, Chicago, 2015.

National Communication Association, Top Panel Award, Instructional Development Division, Las Vegas, 2015.

Purdue University, $3,090.13, Purdue Graduate School Summer Research Grant, 2015.

Purdue University, $100 Communication Graduate Student Association Travel Grant, 2015.

Purdue University, $350 Brian Lamb School of Communication Travel Bursary, 2014.

Partner University Fund Fellowship, $1, 800 The French American Cultural Exchange, 2014.

Purdue University, $350 Brian Lamb School of Communication Travel Bursary, 2013.

Purdue University, $150 Communication Graduate Student Association Travel Grant, 2013.

John Culkin Award for Outstanding Praxis in the Field of Media Ecology, Media Ecology Association, *Figure/Ground Communication*, Managing Editor, 2013.

Purdue University, $250 Brian Lamb School of Communication Travel Bursary, 2012.

Purdue University, $750 Graduate Student Government Professional Grant, 2013.
Purdue University, $122, 808 Tuition Waiver, 2012-2016.

Ryerson University, $500 Communication and Culture Travel Grant, 2008.

Ryerson University, $300 School of Graduate Studies Travel Funding, 2009.

Ryerson University, $12, 000 Ryerson Graduate Award, 2008-2010.

University of Western Ontario, $32, 000 Scholarship, 2008-2010, declined.

Wilfred Laurier University, $12, 000 Scholarship, 2008-2009, declined.

York University, $20, 670 Scholarship, 2008-2010, declined.

Trent University, Dean's Honour Roll, 2007.

Trent University, Form and Context in Literature Award, Satirical Poetry, 2004.

Conference and Workshop Presentations

Iliadis, A. (2015). Nonparticipation and global digital citizenship. 5th Annual International Symposium on Digital Ethics, Loyola University, Chicago. ***Top Student Paper Award***.

Iliadis, A., (2015). Re-ontologizing the body: Data and wearable personal status monitoring (PSM) devices. Wear Me: Art | Technology | Body, University of Ontario Institute of Technology, Oshawa.

Iliadis, A. (2015). Typologies of media non-users. National Communication Association Convention, Las Vegas.

Iliadis, A., Kisselburgh, L., Zoltowski, C., Beever, J., Hess, J., Brightman, A. (2015). Analyzing the development of online ethical reasoning by students in an engineering ethics MOOC using "levels of abstraction" coding of student meta-reflections. National Communication Association Convention, Las Vegas. ***Top Panel Award***.

Iliadis, A. (2015). The right to online nonparticipation. National Communication Association Convention, Las Vegas.

Iliadis, A. (2015). The concept of information in contemporary science (Royaumont, 1962). Philosophy of Communication Conference, Duquesne University.

Iliadis, A. (2015). Socialbots, data, and the role of distributed morality in multi-agent systems. Canadian Communication Association, University of Ottawa.

Iliadis, A., Renner, M. (2015). The professoriate talks: Academic freedom goes online. International Congress of Qualitative Inquiry, University of Illinois at Urbana-Champaign.

Iliadis, A. (2015). The curator, the laggard, and the non-believer: A typology of voluntary information and communication technology non-users. Association of American Geographers, Chicago.

Iliadis, A. (2014). Infostructures: The historical epistemology of a concept. National Communication Association Convention, Chicago.

Hess, J. L., Beever, J., Brightman, A., **Iliadis, A**., Zoltowski, C., Kisselburgh, L., Krane, M. (2014). Deepwater horizon oil spill: An ethics case study in environmental engineering. Society for Ethics Across the Curriculum, Arizona State University, Scottsdale.

Hess, J. L., Beever, J., **Iliadis, A**., Kisselburgh, L., Zoltowski, C., Brightman, A. (2014). A transfer case assessment tool for measuring ethical reasoning abilities of engineering students using reflexive principlism approach. ACM/IEEE (International Society for Electrical Engineering) Frontiers in Education Conference, Madrid.

Kisselburgh, L., Zoltowski, C., Beever, J., Hess, J. L., **Iliadis, A**., Krane, M., Brightman, A. (2014). Effectively engaging engineers in ethical reasoning about emerging technologies: A cyber-enabled framework of scaffolded, integrated, and reflexive analysis of cases. American Society for Engineering Education (ASEE), Indianapolis.

Iliadis, A. (2014). Encyclopedism is a humanism: Double universality and quadruple mediation. Media Ecology Association, Ryerson University, Toronto.

Iliadis, A. (2014). Techné media: Latour and simondon in the garden of technics. Media Ecology Association, Ryerson University, Toronto.

Iliadis, A. (2014). Gilbert Simondon: The concretization of media ecology. Media Ecology Association, Ryerson University, Toronto.

Kisselburgh, L., Zoltowski, C., Beever, J., Hess, J., **Iliadis, A**., Krane, M., Brightman, A. (2013). SIRA framework to develop moral reasoning in engineers. Poster, National Science Foundation Workshop for Ethics and Education in Science and Engineering, Washington, DC.

Iliadis, A. (2013). Individuation in Deleuze, Simondon, and Aristotle. National Communication Association Convention, Washington DC.

Iliadis, A. (2013). Individuation and infostructure: The aesthetics of applicative thinking. Apps and Affect, University of Western Ontario, London.

Iliadis, A. (2013). An abstract machinist: Simondon’s informational artifacts. Society for Literature, Science, and the Arts, University of Notre Dame, South Bend.

Iliadis, A. (2013). What is information artifact ontology? Fifth Workshop of the Philosophy of Information, University of Hertfordshire, Hatfield.

Iliadis, A. (2012). The digital data arms race: Google, Facebook, and Martin Heidegger? Critical Themes in Media Studies, The New School, New York.

Iliadis, A. (2010). Archives of repression: Derrida's psychoanalytic McLuhanism. Marshall McLuhan in a Post Modern World: Is the Medium the Message?, University of Winnipeg, Winnipeg.

Iliadis, A. (2010). Repressive relationality: Towards a critical political economy of contemporary information and communication technology networks. The Social Life of Methods, St. Hugh's College, Oxford University, Oxford.

Iliadis, A. (2010). Finding yourself here: First year student's inquiry into the university. Society for Teaching and Learning in Higher Education, Ryerson University/The Ontario College of Art and Design, Toronto.

Iliadis, A. (2010). A nation talking to itself: Psychoanalytic critiques of contemporary photojournalism in the Middle East. (Re)making (Re)presentation, City University of New York, New York.

Iliadis, A. (2009). Dereferentialization: Universities and the networked society. Activism and the Academy, York University, Toronto.

Iliadis, A. (2009). Back to the symbiotic struggle! How to save cultural studies with French continental philosophy. Intersections, Ryerson and York Universities, Toronto.

Teaching Experience and Field Work

Courses Taught – Instructor

PHIL 490/CS 490/COM 496: Ethics and Philosophy of Information, Purdue University, Department of Philosophy, Brian Lamb School of Communication, Department of Computer Science [x1]

COM 435: Communication and Emerging Technology, Purdue University, Brian Lamb School of Communication. [x1]

COM 336: Advertizing in the Electronic Mass Media, Purdue University, Brian Lamb School of Communication. [x1]

COM 251: Communication, Information, and Society, Purdue University, Brian Lamb School of Communication. [x1]

COM 114: Fundamentals of Speech Communication, Purdue University, Brian Lamb School of Communication. [x2]

Courses Taught – Teaching Assistant

COM 250: Mass Communication and Society, Purdue University, Brian Lamb School of Communication. [x1]

BME 595: Solving Ethical Problems in Engineering: A Course in Multidisciplinary Engineering Ethics, Purdue University, Weldon School, Biomedical Engineering. [x4]

COM 312: Rhetoric in the Western World, Purdue University, Brian Lamb School of Communication. [x2]

SSH 100: Social Scientific Inquiry, Ryerson University, Department of Sociology. [x1]

ENG 108: The Nature of Narrative I, Department of English, Ryerson University. [x1]

Research Assistantships

National Science Foundation (Award Number: 0939370): Center for Science of Information, 2015. https://www.soihub.org/

National Science Foundation (Award Number: 1237868): SIRA Modules for Effectively Engaging Engineers in Ethical Reasoning About Emerging Technologies, 2013-2015. https://engineering.purdue.edu/BME/PRIMEEthics

Book Research, Stuart J. Murray, Ryerson University, Department of English, 2010.

Communication Sector Field Work

Queen Video, Database Consultant, Clerk, 2008-2011.

Trent Radio, Program Director, Archivist, Production Assistant, 2006-2007.

Rogers Communications, Technical Service Consultant, 2004-2005.

Toronto Public Library Morningside Branch, Clerk and Page, 2002-2003.

Guest Lectures

Iliadis, A. (2016). Distributed morality in multi-agent systems. PHI 6938: Digital Ethics, University of Central Florida, Department of Philosophy.

Iliadis, A. (2015). HTML and web publishing platforms. COM 435: Communication and Emerging Technology, Purdue University. Brian Lamb School of Communication.

Iliadis, A. (2015). Building an effective professional web presence in the academy. Communication Graduate Student Association 2015 Spring Colloquium, Purdue University, Brian Lamb School of Communication.

Iliadis, A. (2014). Metadata and digital footprints. COM 325: Interviewing Principles and Practices, Purdue University, Brian Lamb School of Communication.

Iliadis, A. (2014). Qualitative research methods using ATLAS.ti. COM 585: Qualitative Methods for Communication Research, Purdue University, Brian Lamb School of Communication.

Iliadis, A. (2014). Digital politics. COM 251: Communication, Information, and Society, Purdue University, Brian Lamb School of Communication.

Iliadis, A. (2014). Uprisings and citizen participation. COM 251: Communication, Information, and Society, Purdue University, Brian Lamb School of Communication.

Iliadis, A. (2014). Merged, incompatible IT cultures. COM 324: Introduction to Organizational Communication, Purdue University, Brian Lamb School of Communication.

Iliadis, A. (2013). Postmodernism and postmodernity. COM 312: Rhetoric in the Western World, Purdue University, Brian Lamb School of Communication.

Iliadis, A. (2013). Design and technological objects. PHIL 580: Philosophy of Technology, Purdue University, Department of Philosophy.

Academic Service

University

Legislative and Strategic Planning Committee Member, Purdue Graduate Student Government, 2014-2015

Vice Chair, Student Affairs Committee, Purdue Graduate Student Government, 2012-2013

Public Information Committee Member, Purdue Graduate Student Government, 2012-2013
Academic and Professional Development Committee Member, Purdue Graduate Student Government, 2012-2013

Special Projects, Academic and Professional Development Subcommittee Member, Purdue Graduate Student Government, 2013

Purdue Outstanding Graduate Faculty Mentor Award Committee Member, 2013
Purdue Graduate Student Appreciation Week Committee Member, 2013

Organizing Team, Ryerson University Library and Archives Silent Zone Pilot Project, Ryerson University, 2010

Organizer and Adjudicator, Intersections Conference Committee, Ryerson University and York University, 2010

Graduate Representative to the Ryerson University Library and Archives Student Advisory Committee, 2009-2010

Department

Senator, Philosophy Department Representative, Purdue Graduate Student Government, Purdue University, 2014-2015

PhD Member-at-Large, Communication Graduate Student Association, Purdue University, 2012-2013

Co-Vice President, Ryerson University Communication and Culture Graduate Students' Association, 2009-2010

MA Representative to the Ryerson University Communication and Culture Program Executive, 2009-2010

Conferences, Workshops, Symposiums

Program Committee, Seventh Workshop on the Philosophy of Information. "Conceptual Challenges of Data in Science and Technology." University College London, 2015

Panel Chair, "Information and Government." Information Ethics Roundtable 2015, University of Wisconsin-Madison, 2015

Panel Chair, "Critical Data Studies." Canadian Communication Association Annual Conference, University of Ottawa, 2015

Panel Organizer, "Critical Data Studies." Canadian Communication Association Annual Conference, University of Ottawa, 2015

Conference Organizer, Emmanual Levinas Across the Generations and Continents, conference at Purdue University, 2015

Conference Organizer, Communication Graduate Student Association Conference, Purdue University, 2014-2015
Conference Organizer and Adjudicator, Communication Graduate Student Association Conference on Communication Research, Purdue University, 2012-2013

Panel Chair, "Imaginative Historiographies." Intersections, Ryerson University, 2010

Web Development
French Philosophy and Analytic Philosophy in the 20th Century, Purdue University, 2014-2016.

Emmanuel Levinas across the Generations and Continents, Purdue University, 2015.

Organization Communication Mini Conference, Purdue University, 2014.

Communication Graduate Student Association Conference on Communication Research, Purdue University, 2012-2013.

Reviews (years withheld to protect anonymity)

Journal of Communications Media Studies

American Society for Engineering Education, Engineering Ethics Division

Information

Big Data & Society

National Communication Association, Rhetorical and Communication Theory Division

National Communication Association, Great Ideas for Teaching Students
National Communication Association, Philosophy of Communication Division

Canadian Communication Association, Technology and Emerging Media Interest Group

Society for the Philosophy of Information

Deleuze Studies

Parrhesia

International Journal of Knowledge and Systems Science

National Communication Association, Student Section

National Communication Association, Rhetorical and Communication Theory Division

Activities and Projects

Researcher, Centre internationaldes études simondoniennes, 2015-present. http://www.mshparisnord.fr/cides/

Activity Leader, Editorial Projects, Society for the Philosophy of Information, 2014-present. http://www.socphilinfo.org/

Council Member (3 year term), H-Net: Humanities and Social Sciences Online, 2014-2016. https://networks.h-net.org/

H-Rhetor Editor: H-Net: Humanities and Social Sciences Online: H-Rhetor, H-Net's Network on the history of rhetoric, writing, and communication, 2015-present. https://networks.h-net.org/h-rhetor

H-DigiRhet Editor: H-Net: Humanities and Social Sciences Online: H-DigiRhet, H-Net's Network on digital writing, rhetoric, and composition, 2014-present. https://networks.h-net.org/h-digirhet

Head Blogger, *Philosophy of Information and Communication*. http://philosophyofinformationandcommunication.wordpress.com/

Managing and Scholarly Editor, *Figure/Ground Communication*. www.figureground.org

Contributor, Institute for Ethics and Emerging Technologies, 2013-present. http://ieet.org/

Editor, Academia.edu Editor Program. 2015
Contributor, Figure/Ground Communication, Technology and Media Studies, http://figureground.ca/, 2012-2013

York University Free Press Contributor, York University, 2010

Trent University Press Contributor, Trent University, 2007

Trent University Audio Production Workshop, "SOS." Digital Mix, 8 mins; "Three Thoughtful Readers." Multimedia, 2006

Trent University Film Workshop, "For Dziga Vertov." Super 8mm film, 2½ mins; "Malice." 16mm film, 2½ mins, 2006

Trent Radio Board of Directors, 2006-2007

Muted Magazine, Contributor, 2009
Radio Show Host, The Underground Express, Trent Radio 92.7FM, 2003-2007

Only Angels Have Wings, Editor, Staff Writer, 2004-2006

The Underground Express, Head Blogger and Reviewer, 2005-2006

Affiliations

Social Science Research Network, 2015

International Association for Internet and Digital Ethics, 2015

Internet Technology Policy Community, IEEE, 2015

Internet Society, 2015

Modern Language Association, 2015

Canadian Communication Association, 2015

Internet Social Forum, 2015

Conceptual Approaches to Science, Technology, and Innovation, 2014-2015

Special Interest Group, Computers, Information, and Society, 2014-2015

National Center for Faculty Development and Diversity, 2014-2015

Commission for the History and Philosophy of Computing, 2013-2015

Society for Literature, Science, and the Arts, 2013-2014

Society for the Philosophy of Information, 2013-2015

Deleuze Studies Network, 2012-2015
National Communication Association, 2012-2015

Languages and Literatures Circles of Manitoba and North Dakota, 2008-2010

Canadian Association of Cultural Studies, 2008-2010

Association for Cultural Studies, 2008-2010

Society for Teaching and learning in Higher Education, 2008-2010

Languages

South Slavic Languages, Oral

French, Written

Academic References

Daniel W. Smith
Professor
Department of Philosophy
Purdue University
(765) 494-4284
smith132@purdue.edu

Ashley R. Kelly
Assistant Professor
Department of English
Language and Literature
University of Waterloo
(519) 888-4567 x39135
arkelly@uwaterloo.ca

Sorin A. Matei
Professor
Brian Lamb School of
Communication
Purdue University
(765) 494-3429
smatei@purdue.edu

Daniel R. Kelly
Associate Professor
Department of Philosophy
Purdue University
(765) 494-4290
drkelly@purdue.edu

Andrew O. Brightman
Assistant Head,
Biomedical Engineering
Associate Professor,
Engineering Practice
Weldon School of
Biomedical Engineering
Purdue University
(765) 496-3537
aob@purdue.edu

Isabel Pedersen
Canada Research Chair in
Digital Life, Media, and
Culture
Associate Professor
Communication and Digital
Media Studies
University of Ontario
Institute of Technology
(905) 721-8668 x5874
Isabel.pedersen@uoit.ca

PUBLICATIONS

PUBLICATIONS

Brightman, A., Beaver, J., Hess, J., **Iliadis, A.**, Kisselburgh, L., Krane, M., Loui, M., Zoltowski, C. (2016). PRIME ethics: Purdue's reflective & interactive modules for engineering ethics. *Infusing ethics into the development of engineers: Exemplary education activities and programs*. Infusing Ethics Selection Committee, Center for Engineering Ethics and Society, National Academy of Engineering. Washington, DC: The National Academic Press. http://goo.gl/rmMOKi

Iliadis, A. (2015). The right to nonparticipation for global digital citizenship. *International Review of Information Ethics*, *23*, 20-34. http://goo.gl/FBDE0n

Iliadis, A. (2015). Big data. *Communication Booknotes Quarterly*, *46*(2), 54-57. http://goo.gl/xyPMx8

Iliadis, A. (2015). A quick history of the philosophy of information. In Phyllis Illari (Ed.), *The philosophy of information: A simple introduction (pp. 9-27)*. London: Society for the Philosophy of Information. http://goo.gl/eV1iR7

Iliadis, A. (2015). Mechanology: Machine typologies and the birth of philosophy of technology in France (1932-1958). *Systema*, *3*(1), 119-130. http://goo.gl/E4jEeq

Iliadis, A. (2015). Media effects research: A basic overview. *Canadian Journal of Communication*, *40*(3), 582-584. http://goo.gl/SPMq5p

Iliadis, A. (2015). *Why philosophize?* by Jean-François Lyotard. *Dialogue: Canadian Philosophical Review*, online first (May 25), 1-2. http://goo.gl/Q8EllD

Barthélémy, J-H and **Iliadis, A.** (2015). Gilbert Simondon and the philosophy of information: An interview with Jean-Hugues Barthélémy. *Journal of French and Francophone Philosophy*, *23*(1), 102-112. http://goo.gl/OHZbxh

Iliadis, A. (2015). The concept of information in contemporary science (Royaumont, 1962). *Philosophy & Technology*, online first (June 30), 1-3. http://goo.gl/nEXGwh

Iliadis, A. (2015). Two examples of concretization. *Platform: Journal of Media and Communication*, *6*(1), 86-95. https://goo.gl/V0Bmb5
*Spanish. Bernabé Ferreyra (Trans.) https://goo.gl/RbFDCo

Hess, J., Beever, J., **Iliadis, A.**, Kisselburgh, L., Zoltowski, C., Krane, M., Brightman, A. (2014). An ethics transfer case assessment tool for measuring ethical reasoning abilities of engineering students using reflexive principlism approach. *Frontiers in education conference proceedings* (pp. 1-5). Madrid: IEEE. http://goo.gl/fbYxCa

Kisselburgh, L., Zoltowski, C., Beever, J., Hess, J., **Iliadis, A.**, Krane, M., Brightman, A. (2014). Effectively engaging engineers in ethical reasoning about emerging technologies: A cyber-enabled framework of scaffolded, integrated, and reflexive analysis of cases. *American society for engineering education conference proceedings* (pp. 1-17). Indianapolis: ASEE. http://goo.gl/YuXDk6

Iliadis, A. (2013). *The uprising: On poetry and finance* by Franco 'Bifo' Berardi. *Marx & Philosophy Review of Books*, online first (July 30). http://goo.gl/dDxY4A

Iliadis, A. (2013). A new individuation: Deleuze's Simondon connection. *MediaTropes*, *4*(1), 83-100. http://goo.gl/rHExEf

Iliadis, A. (2013). Informational ontology: The meaning of Gilbert Simondon's concept of individuation. *communication +1*, *2*(1), 1-19. http://goo.gl/PmwkAw

Iliadis, A. (2010). The 'claws of absolute necessity': Deleuze on culture. *eTopia*, on-line initiative of *TOPIA: Canadian Journal of Cultural Studies*. http://goo.gl/nwCRAJ