Zhu Li. An Analysis of Two Ontologies of Philosophy. A Master's Paper for the M.S. in L.S degree. April, 2021. 43 pages. Advisor: Ryan Shaw

Unlike many domains that have benefited greatly from the application of linked data, the discipline of philosophy has not fully taken advantage of this relatively new set of methodologies and technologies. Although several researchers have made remarkable attempts to link philosophical resources to assist academic pursuits, the use of their applications have been limited. This paper compares two existing ontologies of philosophy: the PhiloSURFical ontology and the InPhO ontology. It examines particular challenges of developing an ontology of "philosophical ideas" and explores the potential in representing philosophical resources with linked data in order to help with researching philosophical topics.

Headings:

Linked data (Semantic Web)

Ontologies (Information retrieval)

RDF (Document markup language)

Knowledge representation (Information theory)

AN ANALYSIS OF TWO ONTOLOGIES OF PHILOSOPHY

by
Zhu Li

A Master's paper submitted to the faculty
of the School of Information and Library Science
of the University of North Carolina at Chapel Hill
in partial fulfillment of the requirements
for the degree of Master of Science in
Library Science.

Chapel Hill, North Carolina

April 2021

Approved by

_____

Ryan Shaw

Table of Contents

Introduction

Since the start of the 21st century, the amount of research in the humanities that makes use of linked data technologies has grown enormously (Barbera, 2013). Despite bringing interesting changes in disciplines such as history, art history, cultural studies, and literature (Sylva, 2018), linked data has yet to influence philosophy in any major way. Particularly, it seems difficult to represent and link philosophical resources on the Semantic Web due to the rather "abstract" nature of the matters discussed in these resources. Multiple researchers have taken different strategies to do so in order to make it easier for people to perform various kinds of operations with philosophical resources. The goal of this paper is to analyze these strategies and discuss the future possibilities of using linked data to navigate the philosophical world.

The paper first introduces the Semantic Web, linked data and ontologies, followed by a brief discussion of the domain of philosophy. In the literature review section, the paper examines a variety of projects in which researchers have designed different tools to facilitate semantic search of academic resources, including philosophical resources. Then, the paper focuses on two of these projects, PhiloSURFical and InPhO, which have resulted in two ontologies of philosophy. The paper compares these two ontologies in regard to their definition, categorization and relationships of "philosophical ideas" and analyzes their assumptions and why they take different routes to the design of the ontology. Finally, the paper summarizes the results and offers suggestions for future research and projects.

Background

The Semantic Web and Linked Data

The World Web Consortium (W3C) is in the process of realizing its vision of the Semantic Web, a "Web of data" in addition to the classic "Web of documents" (W3C, Semantic Web, n.d.). The Semantic Web is an extension of the World Wide Web "in which information is given well-defined meaning, better enabling computers and people to work in cooperation" (Berners-Lee et al., 2001, p.37). Information can be referred to as linked data as it gets published on the Semantic Web with "well-defined meaning" in a standard format that enables information to be consumed by machines as well as humans. The term linked data (LD) is also used to refer to a method or a set of practices for publishing linked data.

One of the main technologies associated with the Semantic Web is the Resource Description Framework (RDF), a data model designed for representing information in the Web. RDF relies on triples, each consisting of a subject, a predicate, and an object. Subjects and objects can denote almost anything in the world—these things are called "resources" or "entities"—including physical and digital objects as well as abstract ideas or concepts; the predicate in an RDF triple denotes a property, a binary relation that indicates some relationship between the subject and object (W3C, 2014).

Tim Berners-Lee (2009) has coined four principles for publishing linked data, which can be simplified into the following:

1. Use URIs as names for things.
2. Use HTTP URIs so that people can look up those names.
3. When someone looks up a URI, provide useful information.
4. Include links to other URIs. so that they can discover more things.

Ontologies

One of the purposes of the Semantic Web is to make it easier to search for information on the Web. The idea is to let computers integrate RDF from many sources online so that the user can find information they need without going to multiple sources. Traditionally, searching often relies on the presence of search terms in the searched documents, or predictions of information retrieval algorithms. However, it often takes a process of trial and error for the user to come up with effective search terms and results can still turn out unsatisfactory. On the Semantic Web, in order to "link" RDF data from various sources, a common vocabulary is needed, as "vocabularies define the concepts and relationships (also referred to as 'terms') used to describe and represent an area of concern" (W3C, "Ontologies", n.d.). When data in a certain area share a vocabulary, same terms are used to refer to similar entities and relationships and certain constraints are agreed upon in using these terms, and data can be semantically linked. Depending on the requirements of an application, the complexity of the vocabulary varies. The word "ontology" is often used to refer to more complex and formal vocabularies for data integration (W3C, "Ontologies", n.d.).

In Computation Science, an ontology or a computational ontology has had many definitions. Extending on Thomas R. Gruber's definition, Borst (1997) defines an

ontology as “the formal explicit specification of shared conceptualization” (of a certain domain of knowledge), one of the most well-known definitions. After the initiation of the Semantic Web, ontologies have been defined as “new computational artifacts that can provide computational semantics to web content, allowing programs, in addition to data processing, make inferences about this content” (Marcondes, 2013). This new definition does not conflict with Borst’s; new ontologies have the same purpose of conceptual modeling of knowledge domains.

Because of ontologies’ function of defining terms, classifying entities and aiding information retrieval, many consider ontologies as new Knowledge Organization Systems (KOS) (Zeng, 2008; Padmavathi & Krishnamurthy, 2017). Compared with other KOS such as authority files, subject headings and thesaurus, ontologies have a higher degree of “semantic richness”, or “the number of semantic relations between concepts, universals or particulars that KOSs exhibit” (Biagetti, 2020). Unlike other many other KOSs, ontologies are considered highly relevant today. As Biagetti (2020) put it:

> According to this conceptualization [of the Semantic Web], ontologies have been considered the most suitable tools to go beyond the boundaries of the traditional strategies to find and access information. Their relevance appears in machine-to-machine communication, in the exchange of data among systems and in the possibility of facilitating interoperability across heterogeneous systems.

Philosophy as a Discipline

Philosophy, literally meaning “love of wisdom” in Latin, can be defined in many ways. Broadly, it may be understood as “the rational, abstract, and methodical consideration of reality as a whole or of fundamental dimensions of human existence and experience” (Britannica, n.d.) or “all learning exclusive of technical precepts and practical arts” (Merriam-Webster, n.d.). Some see philosophy as a “meta-discipline” and

all other disciplines that entail “concepts, abstraction or reasoning” are sub-disciplines of philosophy (Dean, 2010). In academic institutions, philosophy often resides in the humanities. As those who study philosophy deal with some of humanity’s most general and fundamental questions, philosophical theories are often used as frameworks of enquiry in other humanitarian disciplines. As Blackburn (n.d.) put it, “theory is a record of attempts to understand aspects of ourselves that seem essentially human, such as language, art, history, and literature”.

Because of philosophy’s close relationship to “human existence”, it is challenging for machines to take over the task of information integration and extraction in the domain. The abstract or subtle language of philosophy may not be understood by machines easily. Moreover, philosophy features diverse and contested terminology and contradictory theories, which makes the machine-readable presentation of these complex ideas even harder.

Literature Review

Although numerous ontologies have been created to organize and present various kinds of information, not much work has been done to build ontologies of philosophy in order to search and access information of the discipline. The first related project that entails building semantic relations between abstract ideas of a similar nature to philosophical ideas is the ScholOnto project[1] (Shum et al. 2000), initiated when the Semantic Web was still in its infancy. During the three-year project (2001-2004), a series of tools were developed based on an ontology to overlay a semantic network of scholarly claims on scholarly documents: ClaiMaker to formulate and publish claims, ClaimSpotter for bookmarking and annotating documents on the Web, ClaiMapper for concept mapping, ClaimFinder for searching and filtering, and ClaimBlogger for "semantic blogging" as a new paradigm for scholarly communication. As "any attempt to impose a 'master ontology' on a research field is unlikely to succeed" (Shum et al, 2000), the ontology targets an academic's approach of argumentation instead of knowledge. Assuming that scholars of every discipline are always making a set of claims of the world, the ontology was designed to support scholars by articulating relationships between concepts regardless of the content of their claims. It goes beyond the scope of this paper to verify the assumption that scholars in different discipline make arguments in a similar way that can be captured by one ontology. To my knowledge, the ScholOnto tools were not used widely enough to generate evidence for or against the assumption.

In terms of application, Shum et al. (2000, p.245) were clear about the biggest difficulty for their vision - “the envisaged beneficiaries of the system simply do not have the motivation, skill or time to invest in codifying shared resources to build a critical mass of useful material”. They made efforts to reduce the problem, but the suite of tools still did not gain enough traction. Although the ScholOnto project focused more on supporting the sensemaking process through which scholars make claims than how to link documents in a way to better support navigation of the resources, the rationale behind their ontology applies to any ontology designed to represent a complex field of conflicting views: “The goal is to design a relatively small set of uncontroversial conceptual and relational types which are simple enough to understand without being simplistic, yet expressive enough that most researchers can express the key claims made in most documents” (Shum et al., 2000, p.239). The ScholOnto was a pioneering attempt to use an ontology to structure information in academic documents and promote new knowledge production.

From 2002 to 2004, an EU-funded project named Semantic Web Advanced Development for Europe (SWAD-Europe) presented two demonstration applications relevant to utilize the Semantic Web for communication of ideas. The project aimed to support W3C’s Semantic Web initiative in Europe, “providing targeted research, demonstrations and outreach to ensure Semantic Web technologies move into the mainstream of networked computing” (SWAD-Europe, n.d.). The first demonstrator is Semantic Blogging and Bibliographies, designed to add semantic structure to items shared over the blog channels, providing functions of view, navigation and query. The application utilized multiple vocabularies and machine assisted metadata creation, and

one great challenge was to generate rich metadata without increasing the complexity of the user task (Cayzer, 2004). The demonstrator attracted substantial interest from various people and organizations, but semantic blogging did not go much further afterwards.

The second demonstrator, Semantic Portals, is an informational portal which uses a shared ontology to render all the published information. As an example, a directory of UK environmental, wildlife and biodiversity organizations, was developed and organizations were required to provide RDF data to be included in the directory. The portal provides facets described by a hierarchical thesaurus and the user can select concepts from each facet. One of the important lessons learned in this demonstrator is that informal hierarchies of concepts can be more suited to user navigation than formal ontologies anchored in fundamental distinctions in the world (Reynolds et al., 2004).

In 2006, Pasin and Motta built an ontology of philosophy for students to better understand philosophical concepts and explore scholarly resources. They later applied the ontology to a tool named PhiloSurfical^2, which provided navigation of Wittgenstein's Tractatus Logico-Philosophicus. Similar to the ScholOnto ontology, the PhiloSurfical ontology is based on the questionable assumption that, although the objective or subject matter of research changes constantly, the model of presenting research does not. Compared to the persona of the SchoOnto project who makes claims to persuade others, the imagined PhiloSurfical philosopher works through raising questions and providing answers (Pasin & Motta, 2006). Pasin and Motta tried to not take any philosophical position after they indicated that the ontology is built on the premise of a constructivist epistemology that different philosophical viewpoints consist of interrelated conceptual entities.

Interestingly, around the same time, another project aiming to represent semantic relations among philosophical ideas was carried out by Niepert et al. (2007). As the project focused on building and maintaining a "dynamic formal ontology" in order to extract and manage meta-content of the Stanford Encyclopedia of Philosophy (SEP)[3], it was dubbed the Philosophy Ontology Project (InPhO)[4] (Niepert et al., 2007). The SEP is an open access "dynamic reference work" covering all areas of philosophy, a popular resource for scholars in the humanities. As each entry in the SEP is maintained and kept up to date by an expert or group of experts in the field (About the Stanford Encyclopedia of Philosophy, n.d.), the SEP's meta-content must be updated in a timely manner. The InPhO's ontology was built to support the automated and semiautomated management of the encyclopedia's metadata (Niepert et al, 2007).

A recent application named Relata[5] also identifies complex semantic relationships among humanistic scholarly works. Based on research and design by Rodrigo Ochigame, the project of building Relata started in 2018. As a search tool, Relata aims to facilitate academic searching by "identifying analytical moves or relations—namely, absence, critique, extension, incorporation, reanalysis, and refinement—among scholarly works sharing" (Society for Cultural Anthropology, n.d.). Relata intends to use the "wisdom of the crowds" to populate its database of works – users of Relata are invited to add new works and identify new relations among works. As documentation of Relata is scarce at the moment, it is unclear what advantages it has compared to earlier applications, but its existence does show that the need for semantic search is still unfulfilled.

Comparing Two Ontologies of Philosophy

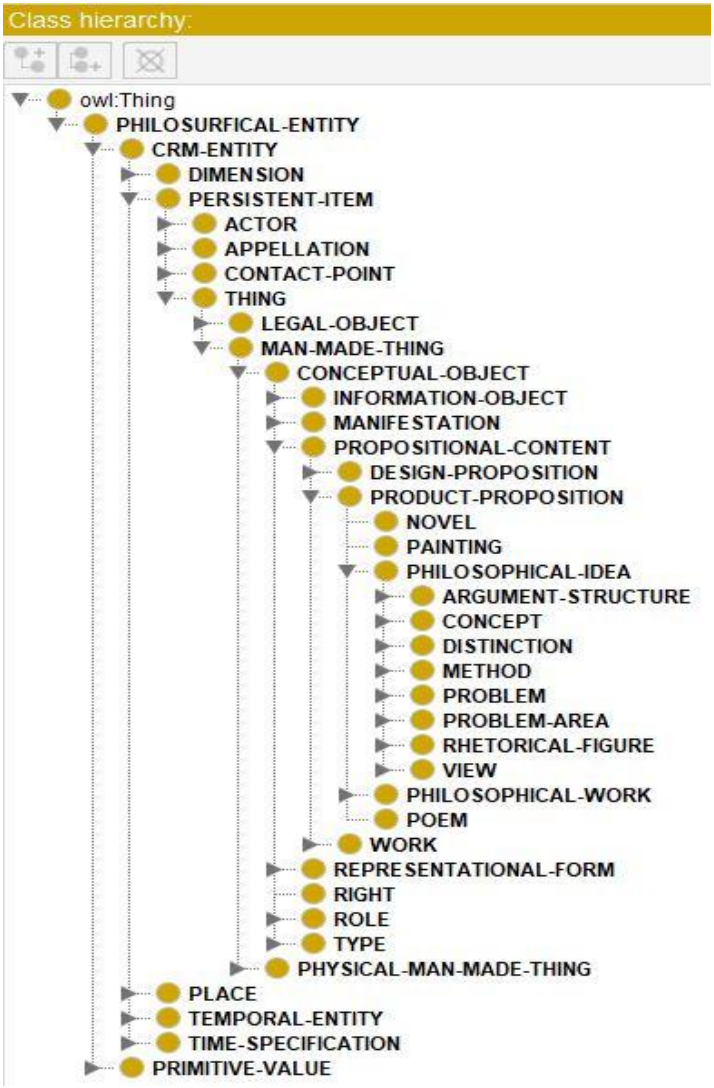
Overview

The PhiloSURFical Ontology

The PhiloSURFical Ontology was made by modifying the CIDOC-Conceptual Reference Model (CRM) to emphasize the history of events related to philosophical ideas through CIDOC-CRM's event-centric approach. The ontology[6], last updated in 2009, has 2533 axioms, 377 classes, and 313 object properties. Similar to the CRM hierarchy, the ontology has the class of CRM-Entity, which is then classified into Dimension, Persist-Item, Place, Temporal-Entity, and Time-Specification (equivalent to Time-Span in the 7.0 version of CIDOC-CRM). As in CIDOC, Conceptual-Object is located in the branch of Persist-Item and contain all abstract entities. PhiloSURFical's Philosophical-Idea class is one of the classes below Product-Proposition, together with Novel, Painting, and Philosophical-Work; Product-Proposition is under Propositional-Content, subsumed by Conceptual-Object.

Figure 1

PhiloSURFical Class Hierarchy



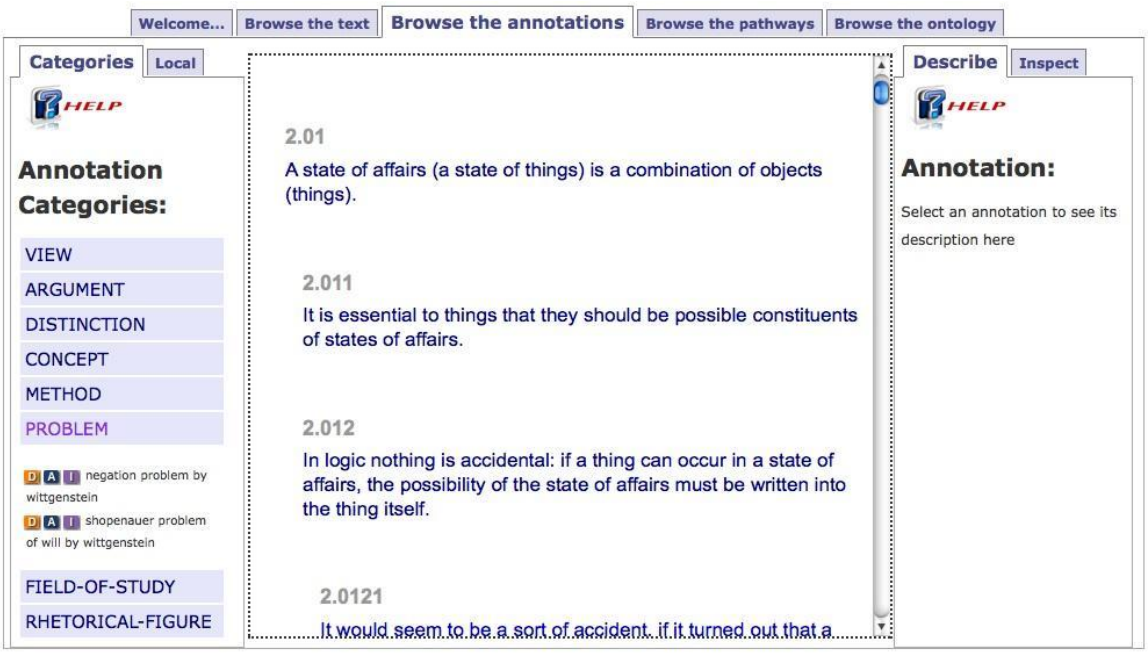
Note. Screenshot of the PhiloSURFical Ontology in Protégé, downloaded from

http://philosurfical.open.ac.uk/onto.html

The PhiloSURFical application was developed to showcase the potential uses of the built ontology. In the application, the ontology was used to generate a semantically enhanced version of Wittgenstein's Tractatus Logico Philosophicus, a classic work in 20th century philosophy. The version contains detailed annotations by a philosophy teacher to serve as an example of a knowledge base built with domain experts based on the ontology. The expert, based on their knowledge, was asked to identify important ideas in the text and assign them to classes in the ontology. As a result, the user of the PhiloSURFical application can browse the text in multiple ways with supplemental information at hand. For example, in the figure below, the left panels show different categories (relevant classes in the ontology) of annotations associated with the text that is currently being browsed, so the user can choose to learn different aspects on different levels of abstraction of the text.

Figure 2

PhiloSURFical Interface

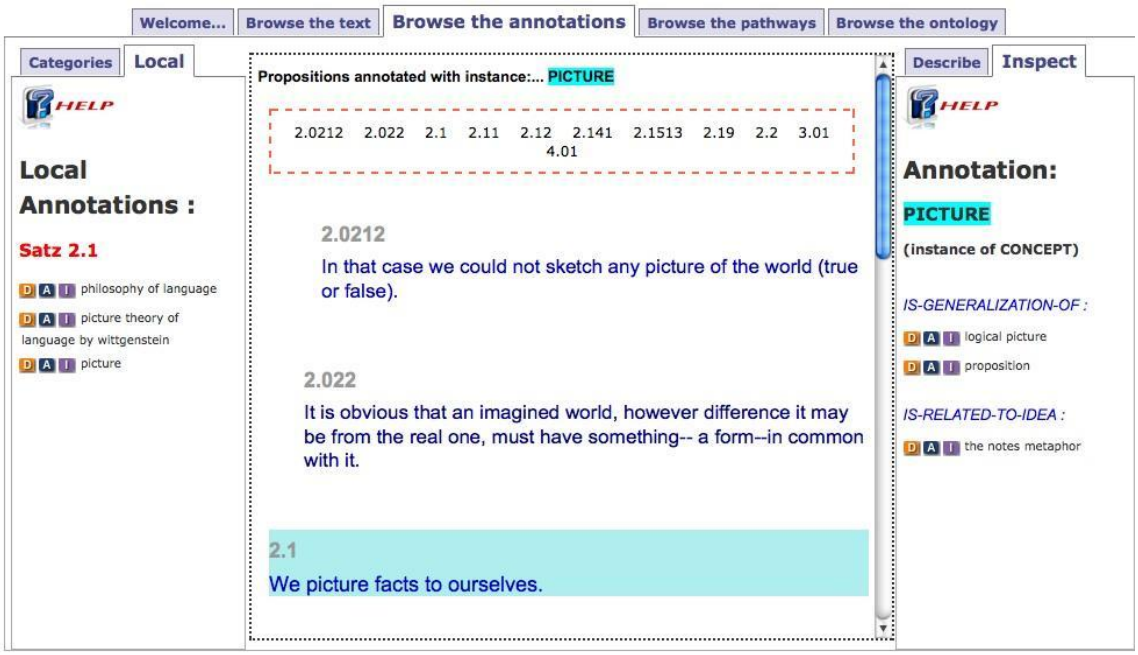


Note. Pasin, M.2007. From PhiloSURFical: Project site. http://philosurfical.open.ac.uk/tour.html

The user can also search for a concept and if the concept is in the annotations, find all the parts related to the concept. The software will show the relationship of the searched concept to other ideas generated by the ontology. The figure below shows what happens when the user searches “picture”. The left side shows the local annotations of the text fragment that the user is looking at and the right side shows related concepts and their relationships.

Figure 3

PhiloSURFical Interface



Note. Pasin, M.2007. From PhiloSURFical: Project site. http://philosurfical.open.ac.uk/tour.html

Different from the relationships in the annotations mentioned before, pathway-triggered relationships go beyond the scope of the Tractatus. The hope was that, with the formalized semantic relations in the ontology and the semantic mappings linking the ontology to other resources on the web, other knowledge bases in the Semantic Web are queried.

The InPhO Ontology

According to the latest available OWL file automatically generated on January 3, 2020 (https://www.inphoproject.org/owl/), the InPhO ontology consists of 19696 axioms, 278 classes and 21 object properties. Many classes exist on the level below “owl: Thing”, but the main categories include Human, Idea, Organization and Publication.

Figure 4

InPhO Class Hierarchy

Note. Screenshot of the 2020-01-03 version InPhO ontology in Protégé. InPhO OWL Files. (n.d.). The InPhO Project. https://www.inphoproject.org/owl/

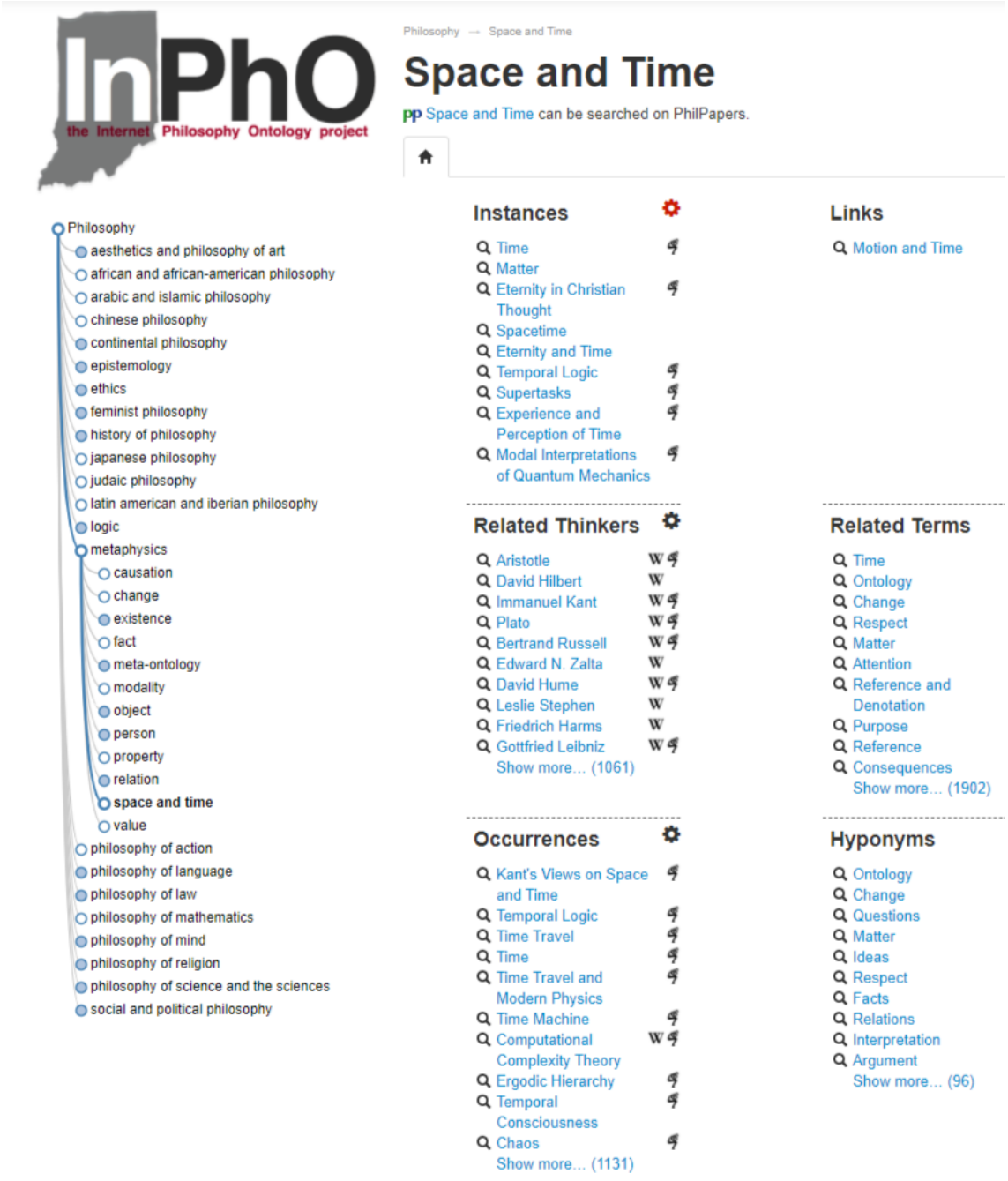
The InPhO ontology can be used in several ways. First, the ontology helps the SEP editors to generate cross references in the articles and the "Related Entries" at the end of each article. Before the InPhO project, SEP's cross-references were built by hand in a time-consuming process with arbitrary guesses. As neither the authors nor the editors have sufficient knowledge of the whole encyclopedia, the author of a new article would be asked to supply a list of keywords with the article, and then editors would repeatedly query the SEP's search engine to find related entries to the article based on the keywords (Niepert et al., 2007). Other problems with the method include inconsistent quality of keywords given by authors as well as the need to update cross-references whenever

articles are updated. As the InPhO ontology provides instances of items related to an article, SEP authors can choose from auto-generated keywords for making cross-reference, which means that cross references can be generated with higher quality and efficiency.

Another application of the ontology is to enable a kind of semantic search and navigation of the covered philosophical resources on the InPhO's website. When the user uses a keyword for searching, the search results will show a series of results related to the closest word in the ontology to the keyword. The page will not only show what online resources you can consult but also related information including, if the keyword points to a topic, instances, links, related thinkers, related terms, occurrences and hyponyms; if the keyword refers to a philosopher, the results will include related thinkers, related terms, influenced by, influenced, teachers, and students. The user can browse the results and find out other important items in philosophy that are probably connected to this term as well as how they are connected. As a result, when the closest term is not accurate enough, the user could possibly go on to find relevant information by choosing a relationship of interest. Meanwhile, a visualization of the word would be on the left side. As shown in figure 5, taking the result of "Space and Time" as an example, the InPhO's interface provides plenty of links for further searching and exploration. The graph on the left side shows the branch where the user is currently located. The user can see the bigger picture while contextualizing the topic of interest in regard to other important areas of philosophy.

Figure 5

InPhO Interface



Note. Screenshot of InPhO webpage. https://www.inphoproject.org/taxonomy/2353

The visualization here is a taxonomy based on the Idea sub-ontology of the InPhO ontology, which allows the user to browse the mainstream landscape of philosophy in the US within a framework of expert level understanding of the domain. One can not only examine an overview of the field but also find subjects of interest by going to an area instead of having to search with a specific keyword.

Instead of organizing ideas by their type, like the PhiloSURFical ontology, the InPhO ontology groups ideas "according to inheritance relationships found in their contents" and creates an Idea subontology that groups ideas according to semantic relevance (Niepert et al, 2007). For instance, philosophy is divided into metaphysics, ethics, logic, philosophy of mind, etc. In this way, the ontology can have more depth and be more intuitive to scholars. By means of statistical text processing, the ontology has taxonomic and non-taxonomic information extracted from both the documents of the SEP and external sources such as Wikipedia and academic genealogy datasets (Niepert et al, 2007); some categories and all the instances on the lowest level are directly from entries of the SEP. This approach resembles the PhilPapers's categorization system[7] for retrieval of philosophical documents; the PhilPapers taxonomy was built by philosophers and is currently maintained and refined by philosophers and PhilPapers users.

One of the great achievements of InPhO is that it uses automated and semi-automated methods for populating and managing the ontology together with domain experts' feedback to ensure the quality of the project. The ontology enabled a series of

new functions such as semantic search, automated generation of cross-references and tables of contents, and ontology-driven conceptual navigation (Niepert et al., 2008). For each instance of the InPhO ontology, the user can explore related content using provided links of “Related Thinkers”, “Related Terms”, “Occurrences”, “Hyponyms”, etc. However, as automated methods were not able to capture enough data to warrant the efforts, InPhO did not populate and use most of their more granular non-taxonomic properties, such as attacked_view, aware_of, commits_to.

Purpose

The PhiloSURFical ontology and the InPhO ontology are different from each other largely because they were designed for different purposes. The PhiloSURFical ontology was mainly built for pedagogical purposes – to help students navigate philosophical literatures, learn philosophy with richer contextual information and “through serendipitous discovery of relevant resources” (Pasin & Motta, 2011). The idea is to generate metadata by asking philosophers to annotate philosophical texts, representing their knowledge by instantiating the ontology.

The InPhO project’s major goal was to develop an information management tool for large-scale and complex digital reference works that are being updated regularly. Specifically, when the “InPhOrmers” started to work on their first collection of texts, the Stanford Encyclopedia of Philosophy (SEP), the ontology was built to generate and manage the SEP’s machine-readable meta-content and support new functions such as semantic search and ontology-driven conceptual navigation of the encyclopedia.

Scope

Although both ontologies were built for the domain of philosophy, their scopes differ from each other. Like the definition of the scope of CIDOC CRM, it is useful to distinguish an ontology's intended scope and practical scope: the intended scope is the domain that the ontology would ideally aim to cover and is expressed as a definition of principle; the practical scope is the current coverage of the ontology and is expressed "in terms of the reference documents and sources that have been used in its elaboration" (CIDOC CRM Special Interest Group, n.d.). Based on these understandings, the intended scopes of the two ontologies appear similar in that both ontologies would ideally cover the whole domain of philosophy, but their practical scopes are different because of the different texts that the ontologies have been based on or applied to.

As the practical scope evolves with new sources, the current practical scope of the InPhO ontology is based on the SEP, Internet Encyclopedia of Philosophy, PhilPapers, HathiTrust/Google Books Collection. The PhiloSURFical ontology has only been used on Wittgenstein's Tractatus Logico-Philosophicus, so the PhiloSURFical ontology's the practical scope is limited by the work.

The scope of an ontology is also determined by its purpose, so the two ontologies provide answers to different types of questions. InPhO's purpose of managing specific materials determined that it covers featured information from its sources. For example, after the PhilPapers had been incorporated, the ontology started providing more information on journals. Compared to the InPhO ontology, the PhiloSURFical ontology

is intended to provide students with various learning pathways (“learning pathway” as a “system of specially stored and organized narrative elements which the computer retrieves and assembles according to some expressed form of narration”); it was not created based on specific resources and aims to be able to represent any old or new resources in Philosophy. As a result, the PhiloSURFical ontology’s scope might include a wider range of perspectives on the resources, such as the historical narrative and the geographical narrative, which are not highlighted in InPhO.

Philosophical Ideas

When building ontologies, the PhiloSURFical team and the InPhO team both focused their efforts on representing “philosophical ideas", but they define these ideas differently.

In the PhiloSURFical ontology, one of the major extensions of CIDOC-CRM is the Philosophical-Idea class, which contain all philosophical ideas. According to the class’s “rdfs: comment” that is used to clarify the meaning of the class, the Philosophical-Idea class contains philosophical ideas that are defined as “a propositional content with a specific importance within the philosophical world” (“philosurfical.owl”). The class is a subclass of the Product-Proposition class, which belongs to the Propositional-Content class. The Propositional-Content class is under the Conceptual-Object class, which is defined in the same way as CIDOC-CRM defines its conceptual-object. In Version 6.2.2 of CIDOC-CRM, “E28 Conceptual Object” “comprises non-material products of our

minds and other human produced data that have become objects of a discourse about their identity, circumstances of creation or historical implication".

The PhiloSURFical ontology's definition of philosophical ideas comes from Riichiro Mizoguchi's theories of ontology, specifically his ontology of "representation". According to Mizoguchi (2004), "representation" means "a proposition coded in a form" such as novels, music and symbols; a representation is different from other objects because it is a content-bearing thing. He presents an ontology of representation and divides a representation into "form", named "representational form", and "content", named "proposition". The latter part refers to "a proposition which the author of the representation would like to convey through the Representation" (Mizoguchi, 2004). He further categorized "representational form" into "symbol sequence", "speech", "still image" and "motion image" and "proposition" into two kinds: "design proposition" and "product proposition".

Design proposition is a specification of the production of something, such as a piece of music played by a musician; Product proposition itself is the product of the representation, such as a novel. In the PhiloSURFical ontology, the Philosophical-Idea class is placed under the Product-Proposition class, together with painting, novel, philosophical-work, and poem, because philosophical ideas are seen as "product propositions, as they are important in themselves, for argumentation or theoretical purposes, and not for specifying an action" (Mizoguchi, 2004).

InPhO's definition of the Idea branch is less clear. As a major class under "owl: Thing", Idea class is separated from Agent (Human and Organization), Publication, Nationality, and Profession. The Idea branch initially was created with a single goal of covering all the keywords of the SEP entries, which "correspond directly to the items of interest in the domain of philosophy" (Niepert et al, 2007). As a result, InPhO's ideas could virtually be defined as philosophical terms that did not refer to humans, organization, publication, etc. and were deemed noteworthy in the discipline of philosophy by the SEP editors at the time. This definition mirrors the PhiloSURFical's definition in which "philosophical ideas" should be important for the discipline and they are differentiated a philosopher, a work of philosophy, or a philosophy school.

Although philosophical ideas are considered crucial in both ontologies, neither team of ontologists found much literature to inform of their endeavors. When Niepert et al. (2007) said little progress has been made towards creating ontologies of ideas, as most were referring to a special kind of ideas, since ontologies are always representing ideas based on its definition that "[a]n ontology is a specification of a conceptualization" (Gruber, 1992). An ontology of chocolate, for example, is an ontology of ideas/description/specification of chocolate. However, an ontology of philosophical ideas is an ontology of ideas of philosophical ideas, and not many ontologies of ideas of ideas have been created.

Among the many obvious differences between chocolate and ideas, the most fundamental is that ideas do not need space to exist and thus appear to be more abstract in

a common sense. The PhiloSURFical's framework, i.e, Mizogochi's theories, may help to explain why building an ontology of abstract philosophical ideas may be more difficult than building other ontologies. According to Mizogochi (2004), philosophical ideas, defined as propositions, need some kind of form to be expressed as representations, such as spoken or written language; then, representations need to be embodied and become a represented thing to be sensible for humans. For example, "[a] sentence 'This is a book' is a representation in the form on natural language (English) and what you see is its printed realization on a sheet of paper which is a represented thing" (Mizogochi, 2004). "As a representation has several deeply related concepts such as its embodiment, the mode or the form of representation and its content…", it may be hard to hard to know what is an instance of a representation (Mizogochi, 2004) and classify representations in a sophisticated way.

Classification of Ideas

In the PhiloSURFical ontology, ideas are categorized according to their functions into eight main types: argument-structure, concept, distinction, method, problem, problem-area, rhetorical-figure, view. These different types of ideas play different roles "in the construction of viewpoints, and, more broadly, having a recognizable function in the process of interaction and succession of viewpoints within the whole history of thought" (Pasin & Motta, 2011, p.244). Taking an entirely different approach, the InPhO ontology does not differentiate ideas by their type, but by their meaning. Following the discipline's tradition, philosophical keywords in the SEP are classified according to

paradigmatic relationships found in their contents rather than by their kinds. Consequently, if the SEP's keywords correspond to all the items of interest in philosophy, the ontology can naturally classify philosophical ideas along the standard sub-specializations of philosophy: metaphysics, epistemology, logic, ethics, and aesthetics. Because of the ontologies' different approaches to classification, the ontologies were built in different ways. The PhiloSURFical ontology only structured types of ideas, so users were able to create labels that can act as access points as they go through documents. The InPhO project started with a hand-built ontology by experts with existing keywords of the SEP as classes or instances, then documents were assigned to these classes or instances.

How philosophical ideas are classified has greatly influenced how documents can be indexed using the two ontologies. As one of the shared goals of the two ontologies is information retrieval, it is worthwhile to compare their indexing with the most used index for retrieving documents– subject indexes. As Lancaster (2003, p.22) pointed out, subject indexing, or subject cataloging, essentially involves the activity of subject classification, "i.e., forming classes of objects on the basis of their subject matter". Subject indexing usually consists of two stages: the conceptual analysis stage, in which subjects are attributed to documents, and the translation stage, in which the indexer represents the product of conceptual analysis by means of a term or terms drawn from some vocabulary and assign resulting labels to documents (Lancaster, p.21). The building of the two ontologies also involves subject classification. However, in the conceptual analysis stage,

the indexer not only needs to determine what a document is about, but also what other attributes it has, such as “which philosopher does it involve?” and “what problem can it solve?” The intention behind this is that documents can be found more easily with a more detailed classification.

One of the most useful attributes that are utilized in these ontologies are philosophical ideas that are not necessarily subject matter of a document. These ideas may be a concept that a philosopher defines or a viewpoint that is cited for illustration. One may argue that the classes and instances in these two ontologies are all subjects, which leads to the question “What is a subject?”. The use of “subject”, a fundamental concept in Library and Information Science (LIS), “is part of the broader use of the concept that refers to all kinds of utterances” (Hjørland, 2017). Based on this definition, in its broadest sense, a subject can be anything that is uttered in a document or one thinks is uttered in a document.

Despite the wide use of subject indexes, interpretation of subject indexing can differ greatly. As elaborated by Furner (2012), there exists “a nominalist-realist spectrum of views about aboutness and subjecthood”: at the nominalist pole, aboutness is considered as a relation between entities and subjects are subjective labels within a system; at the realist pole, aboutness is considered as a property of entities and subjects are designated by documents. This difference of views is mirrored in the distinction of theoretical positions between request-oriented views versus document-oriented views (Hjørland, 2017). Traditional indexing has been content- or document-oriented: people

think that subjects are inherent in documents and good indexers can correctly find the subjects of a document. Request-oriented indexing means that the user's request should determine how documents are indexed and subjects are created by indexers to meet the user's request.

Depending on which position one takes—nominalist or realist, request-oriented or content-oriented—one may have different opinions towards the two ontologies' classes and instances. If we understand subjects in a narrow sense, a document may only have a few subjects, and these ontologies are indexing concepts in addition to their subjects. In 1980, Bernier differentiated subject indexing from indexing of concepts, topic and word. Comparing subjects and concepts, he argued that subjects are what authors are working and reporting on while concepts may just be what authors use during the process; "indexes to concepts that authors et al. have and report … include subjects; but they also include much more than subjects. Concept indexes are bulky; entries to the subjects in them are greatly diluted" (Bernier, 1980, p.192). Bernier (1980) also pointed out that it is easy for subject indexers to drift into indexing concepts and words rather than subjects and illustrated this with an example of chromatography in documents of chemistry. When chromatography was first introduced into chemistry, it was an interesting subject for chemists as they studied its procedures and techniques. However, as chromatography got well studied and became a standard technique of chemical analysis, it was no longer a subject in many studies, but instead became a concept considered and expressed by chemists who were working on other subjects. Surprisingly, many indexers continued to

index the concept of chromatography under the subject heading Chromatography (Bernier, 1980). Although indexers need to resist the indexing of concepts and words when indexing subjects, indexing concepts can be useful sometimes, as in the case of these two ontologies.

In regard to the aspect of subject indexing, the two ontologies have also been built on the aforementioned different assumptions of aboutness and subjecthood as well. Interestingly, as InPhO ontology classifies documents according to the keywords of the SEP and chooses labels from the existing vocabulary of it, the ontology's organization approach resembles traditional subject indexing more. In comparison, the vocabulary that PhiloSURFical can draw labels from is likely to be the entire English vocabulary because of great variation among annotators' conceptual analysis and choice of labels, the ontology appears to allow much more freedom than InPhO; an idea or concept can be indexed using PhiloSURFical but not necessarily with InPhO.

Additionally, InPhO's "subjects" appear to be more content-oriented or document-oriented with the project's automatic indexing process. The reliance on term occurrences is built on the assumption that the machine can predict what documents are talking about, which is treated as an objective fact. As quality indexing depends on being able to clearly define subjects, the effectiveness of the two ontologies can be gauged according to their conceptualization of philosophical ideas.

Relationships

Ontologies' "semantic richness" compared to other KOSs means that they allow a greater number of relationships between terms. These relationships can not only specify class hierarchies but also enable reasoners to make inferences about instances, producing new knowledge. KOSs such as thesaurus (Clarke, 2019) have been able to provide equivalence relations (Use for or Variant), hierarchical relations (Broader Term, Narrower Term) and associative relations (Related Term), but more specific and complex relations between terms can be provided in ontologies as they present properties for each class.

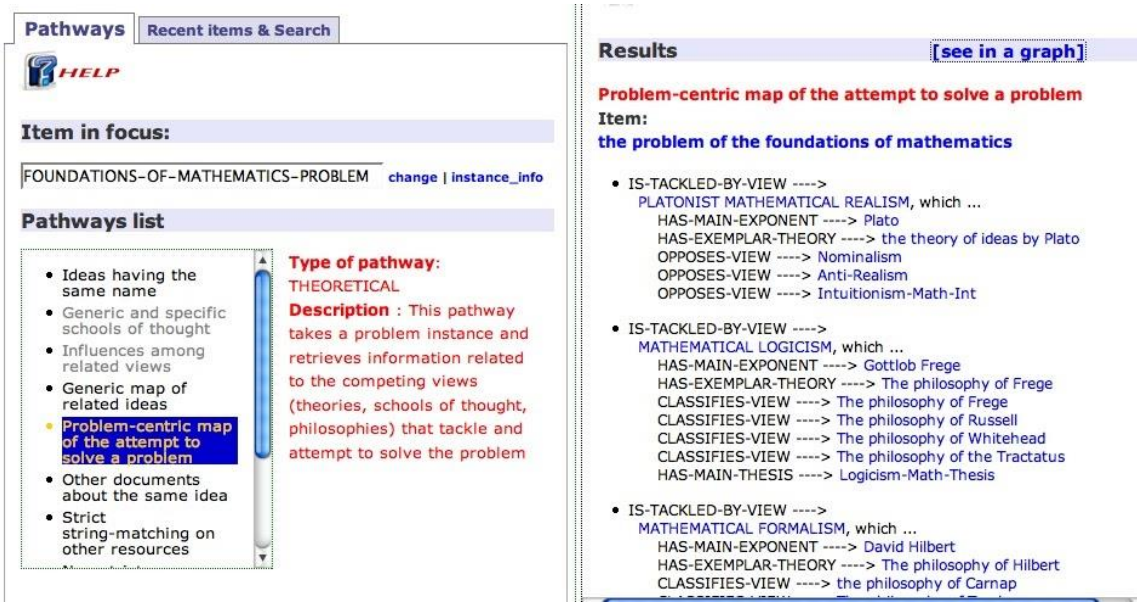
In the PhiloSURFical ontology, as philosophical ideas are organized into eight classes with various different properties, a great many relationships are being defined. With an object property count of 303, the ontology has specific relationships with well-defined domains and ranges. For instance, Concept, a subclass of Philosophical-Idea, has the such as Causes-Concept, Defined-By-View, Has-Opposite-Concept, Has-Related-Concept, Is-Equivalent-To, Is-Generalization-Of, Is-Specialization-Of, Requires-Concept).

PhiloSURFical's relationships between ideas can be used in two major ways. First, when the user browses the annotated text in the PhiloSURFical application, they can see the certain topics' local annotations, in which related categories and the nature of the relationships are provided. By means of these relationships, the user can navigate the text in a non-linear manner. Another way to use the relationships is when the user has a

topic in mind and search for related terms. The application shows all the entities related to the searched item, displayed as a list of triples representing subject-predicate-object relationships, which are obtained through the semantic relations formalized in the ontology or planned releases of semantic mappings linking the ontology to other services available on the web ("Tour"). The application can also present a pathways list—such as "influences among related views", "generic map of related ideas", "problem-centric map of the attempt to solve a problem", essentially further indexing related topics.

Figure 6

PhiloSURFical Interface



Note. Pasin, M.2007. From PhiloSURFical: Project site. http://philosurfical.open.ac.uk/tour.html

The InPhO ontology initially also contained non-taxonomic relationships that are usually not represented in traditional KOSs, but these relationships were not used in the final application. As the InPhO ontology does not differentiate ideas by their functions, the number of relationships is much smaller than PhiloSURFical; only several relationships involving ideas exist between ideas and humans or publications. The population of more complex relations required the development of extra algorithms to find instances of relations between ideas and ideas or philosophers from various sources, but the team did not find enough data to formally define these relations. As a result, the ontology's relationships are not much different from traditional KOSs' equivalence relations, hierarchical, and associative relations.

Table 1

InPhO's Non-taxonomic Relationships between Human, Idea and Publication

	Human	Idea	Publication
Human	criticized, defended, discoursed_with, dissertation_advisor_of, has_influenced,	attacked_view, aware_of, created_view, worked_on	edited, wrote
Idea		commits_to, opposed_to	
Publication		discusses	cites, published_in,

Note. Based on the 2020-01-03 version InPhO ontology in Protégé. InPhO OWL Files. (n.d.). The InPhO Project. https://www.inphoproject.org/owl/

Discussion

Several different features of the two ontologies are determined by their different purposes. For instance, due to the PhiloSURFical ontology' pedagogical purpose, the ontology prioritizes contextualizing philosophical ideas for the purpose of informing student learners. This had made the CIDOC-CRM more relevant in that it enables the ontology to have a great many entities related to events: "the temporal entities regarding events related to the academic life and to the life of philosophers", "the temporal entities related to production and modification of philosophical ideas" and "the temporal entities representing philosophical historical periods" (Pasin & Motta, 2011). This makes it easier for students to conceptualize the history of philosophy. The InPhO ontology was initially built upon the SEP and designed for the SEP's metadata management, so InPhO's hierarchy of concepts was essentially determined by the SEP's texts and contributors.

Due to the different purposes of the two ontologies, the InPhO ontology's target user group is wider than the PhiloSURFical ontology. Initially, when the ontology was only for managing the SEP, any SEP user could take advantage of the ontology. At this stage, the user could be a student of philosophy, a scholar or a student of other disciplines who needs to know more about certain theories, or anyone curious outside academic institutions. As the InPhO project incorporated other philosophical resources online, the user group expanded, as one could use the ontology to access not only more academic papers on topics of philosophy, but also more non-academic resources online. Because of

the diversity of InPhO users, the ontology can be described as more general, constrained within mainstream philosophy standpoints.

The two ontologies both have strengths and weaknesses. The PhiloSURFical ontology, revealing intricate relationships that can be found among philosophical resources, performs better for a user that browses texts for serendipitous discoveries or tries to dig deep in one document. The InPhO ontology, offering a better bird-view of the philosophical world, helps a user better with finding relevant resources to a topic of interest.

Both ontologies can help researchers of philosophy and other humanities that employ philosophical theories, as studies have shown that humanities scholars, especially interdisciplinary ones, use diverse research methods (Brockman et al., 2001; Toms & O'Brien, 2008). Most humanities scholars conduct research using thematic or semantic approaches—they usually examine text or texts surrounding a theme, such as "benevolence in a nineteenth century poet and essayist", "regional trends in Irish-American fiction", "study of domestic assault in Victorian fiction" (Toms & O'Brien, 2008). When a scholar conducts research, InPhO might provide more help when initiating searching or zooming out after some reading to find more relevant resources; PhiloSURFical might prove more useful when examining texts.

PhiloSURFical focuses more on making an ontology of theoretical soundness while InPhO emphasizes the practicality of the ontology. PhiloSURFical, with a clear theoretical base that determines what should, or should not be included in a class, is compatible with the CIDOC CRM, which makes it fit for an existing description of the

real world. However, the PhiloSURFical data model is like a formula for users, which may appear awkward for philosophy students and scholars, who are not used to formulas. In fact, Pasin himself when working on another tool to facilitate humanities research years later, stated:

> almost all humanities scholars spend their time developing their own original interpretation of the materials they study, and aim to explore new concepts and paradigms about them which they present in their articles and books (see Brockman et al 2001 and in Palmer et al 2009). The scholarship does not start out with predefined formal structures, but begins with a set of vague notions and insights in the scholar's mind as they read that only over time emerge clearly enough to be described in published work. (Bradley & Pasin, 2012)

Bradley & Pasin (2012) also concluded that even when humanities scholars' ideas are mature enough for publication, "their 'model' may be only partially compatible with formal ontologies. Moreover, Nieper et al. (2007) also pointed out that an ontology that classifies ideas according to their kinds instead of their content may not only seem "contrived" but also far from the ubiquitous intuitive searching and thinking patterns in the domain of philosophy.

Another important consideration is whether an ontology of philosophy can be put to sustainable use. In this regard, InPhO may be preferable. InPhO uses a combination of automatic methods and human input to gather information in order to reduce the burden on the domain experts. Although using statistical methods with experts may cause certain biases, such as only representing mainstream theories, this approach is beneficial for the quality of information, and it is more likely to motivate experts to work. For instance, a SEP author's workload might decrease because of the automatic generation of cross-references. In the case of PhiloSURFical, however, it is hard to get experts to annotate many long texts. This is a persistent problem for these applications as noted previously.

Conclusion

The paper examines two ontologies of philosophy: the PhiloSURFical ontology and the InPhO ontology. It compares the two ontologies in regard to their purpose, scope, philosophical ideas and relationships, and analyzes their advantages and disadvantages. It contributes to applying linked data technologies to the discipline of philosophy, ontology engineering of philosophical ideas, and developing better tools for philosophy scholars and students as well as people outside academic institutions.

Based on this analysis, one potential future project may be to explore the possibility of combining the two ontologies. The combined ontology might contain more relationships than the InPhO ontology yet still be user-friendly. Depending on the specific context of use, different features of the two ontologies could be adopted.

Another potential project to work on is to find more ways to take advantages of experts and machines to generate more complex semantic relationships than those in traditional knowledge management tools - How do we make it easier for experts to provide detailed semantic information along with their works? How can machines capture more relationships within a given text? Answering these questions would not only benefit the discipline of philosophy but also other disciplines that use philosophical theories. Better ways to structure “abstract ideas” are essential for building the Semantic Web and enabling more domains to benefit from linked data technologies.

Notes

1. While the ScholOnto's tools cannot be used now, documentation of the project can be found here: http://projects.kmi.open.ac.uk/scholonto/software.html
2. PhiloSurfical is no longer being maintained. Visualizations of the tool can be found here: http://philosurfical.open.ac.uk/index.html
3. http://plato.stanford.edu/
4. Now named the Internet Philosophy Ontology (InPhO): https://www.inphoproject.org/
5. https://relata.mit.edu
6. The latest ontology can be found here: http://philosurfical.open.ac.uk/onto.html
7. Information on the PhilPapers Categorization Project: https://philpapers.org/help/categorization.html

References

Barbera, M. (2013). Linked (open) data at web scale: research, social and engineering challenges in the digital humanities. *JLIS.It, 4*(1), 91–101.

Berners-Lee, T. (2009). *Linked Data*. https://www.w3.org/DesignIssues/LinkedData

Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The Semantic Web. *Scientific American*, 284(5), 34–43. https://doi.org/10.1038/scientificamerican0501-34

Biagetti, M. T. (2020). Ontologies (as knowledge organization systems). In B. Hjørland & G. Claudio (Eds), *Encyclopedia of Knowledge Organization*. https://www.isko.org/cyclo/ontologies

Blackburn, P. (n.d.). Theory for the Humanities. http://theory-humanities.weebly.com/

Borst, W. (1997). Construction of Engineering Ontologies. [PhD thesis, University of Twente]. Centre for Telematics and Information Technology (CTIT).

Bradley, J., & Pasin, M. (2012). Annotation and Ontology in most Humanities research: accommodating a more informal interpretation context. *NeDiMaH Ontology Workshop*. http://pliny.cch.kcl.ac.uk/docs/NeDiMaH-proposal.pdf

Cayzer, S. (2004). SWAD-Europe deliverable 12.1.4: Semantic Blogging – Lessons Learnt. https://sciwheel.com/work/item/10942186/resources/11406313/pdf

CIDOC CRM Special Interest Group. (n.d.). Scope. http://www.cidoc-crm.org/scope

Dean, T. (2010). What is Philosophy. https://ockhamsbeard.wordpress.com/2010/06/11/what-is-philosophy/

Dextre Clarke, S. G. (2019). Thesaurus (for Information Retrieval). In B. Hjørland & G. Claudio (Eds), *Encyclopedia of Knowledge Organization*. https://www.isko.org/cyclo/thesaurus

Shum, S. B., Motta, E., & Domingue, J. (2000). ScholOnto: an ontology-based digital library server for research documents and discourse. *International Journal on Digital Libraries*, *3*(3), 237–248. https://doi.org/10.1007/s007990000034

Furner, J. (2012). FRSAD and the Ontology of Subjects of Works. *Cataloging & Classification Quarterly*, *50*(5–7), 494–516. https://doi.org/10.1080/01639374.2012.681269

Hjørland, B. (2017). Subject (of documents). *Knowledge Organization*. 44, no. 1: 55-64.

InPhO OWL Files. (n.d.). The InPhO Project. https://www.inphoproject.org/owl/

Jay, P. (2014). critique and theory in the history of the modern humanities. In R. Bod, T. Weststeijn, & J. Maat (Eds.), *The making of the humanities: volume III: the modern humanities* (pp. 655–666). Amsterdam University Press. https://doi.org/10.1515/9789048518449-044

Marcondes, C. H. (2013). Knowledge organization and representation in digital environments: relations between ontology and knowledge organization. KNOWLEDGE ORGANIZATION, 40(2), 115–122. https://doi.org/10.5771/0943-7444-2013-2-115

Merriam-Webster. (n.d.). Philosophy. In *Merriam-Webster.com dictionary*. https://www.merriam-webster.com/dictionary/philosophy

Mizoguchi, R. (2004). Tutorial on Ontological Engineering: Part 3: Advanced Course of Ontological. *New Generation Computing*. 22(2), 193-220. https://doi.org/10.1007/BF03040960

Niepert, M., Buckner, C., & Allen, C. (2007). A dynamic ontology for a dynamic reference work. *Proceedings of the 2007 Conference on Digital Libraries - JCDL '07*, 288. https://doi.org/10.1145/1255175.1255230

Niepert, M., Buckner, C., & Allen, C. (2008). Answer Set Programming on Expert Feedback to Populate and Extend Dynamic Ontologies. *Twenty-First International FLAIRS Conference*. https://www.aaai.org/Papers/FLAIRS/2008/FLAIRS08-116.pdf

Padmavathi, T., & Krishnamurthy, M. (2017). Semantic web tools and techniques for knowledge organization: an overview. *KNOWLEDGE ORGANIZATION*, 44(4), 273–290. https://doi.org/10.5771/0943-7444-2017-4-273

Pasin, M. (2006). *An ontology for the description of and navigation through philosophical resources* (Enrico Motta, Trans.). European Conference on Computing and Philosophy (ECAP 2006), Trondheim, Norway. http://oro.open.ac.uk/23623/

Pasin, M. (2007). PhiloSURFical: Project site. http://philosurfical.open.ac.uk/tour.html

Pasin, M., & Motta, E. (2011). Ontological requirements for annotation and navigation of philosophical resources. *Synthese*, *182*(2), 235–267. https://doi.org/10.1007/s11229-009-9660-3

Philosophy. (n.d.). Encyclopedia Britannica. https://www.britannica.com/topic/philosophy

Society for Cultural Anthropology. (n.d.). Relata. https://culanth.org/engagements/relata

SWAD-Europe. (n.d.). SWAD-Europe. https://www.w3.org/2001/sw/Europe/

World Wide Web Consortium (W3C). (2014). RDF 1.1 Concepts and Abstract Syntax. https://www.w3.org/TR/rdf11-concepts/

World Wide Web Consortium (W3C). (n.d.). Ontologies. https://www.w3.org/standards/semanticweb/ontology#summary

World Wide Web Consortium (W3C). (n.d.). Semantic Web. https://www.w3.org/standards/semanticweb/

Zeng, M. L. (2008). Knowledge organization systems (KOS). *KNOWLEDGE ORGANIZATION*, *35*(2–3), 160–182. https://doi.org/10.5771/0943-7444-2008-2-3-160