Project Narrative

Significance and Impact

A predominant scholarly problem is how to comprehend the explosion of high-quality scholarship available in digital formats on the Internet. Humanities scholars, like all academics, are increasingly reliant on the World Wide Web for access to scholarly materials and they are rapidly transferring traditional journals and rare archives to digital formats, further exacerbating the problems of information overload. Available search engines have failed to solve the problem of meaningful access, and most users, including students and scholars, lack the necessary skills to construct effective search queries.

In light of these challenges, some digital humanities projects have begun to build and maintain collections using machine-readable and structured representations such as XML and RDF. In recent years, the Linked Data initiative[2] has gained considerable traction. Its goals are to create large and interconnected collections of open and structured data repositories. Examples are (a) VIVO[3] which provides machine readable facts about researcher interests, activities, and accomplishments, enabling the discovery of research and scholarship across disciplines; and (b) DBPedia[4] a data repository that contains structured information extracted from Wikipedia. The primary motivation for this project is that data in isolation does not maximize its potential utility. By including links between entities in a data collection, additional facts can be inferred, making the collection more valuable than the sum of its parts and giving information context and interoperability. This year, there was a first conference on "Networked Humanities: Art History in the Web"[5] where participants discussed how semantic web technology could benefit humanity scholars and merge split collections. Every year, digital humanity projects present their work at the International Conference for Digital Scholarship in the Humanities[6] (DH) and the number of collections is growing steadily. The motivation of our proposal is the desire to accumulate the knowledge of these mostly disconnected collections. Figure 1 depicts a small fragment of the Linked Data cloud where each circle corresponds to a single dataset and arrows indicate the existence of links between objects.

---

[2] http://linkeddata.org

[3] http://vivoweb.org/about

[4] http://dbpedia.org/About

[5] http://www.esf.org/index.php?id=6726

[6] http://dh2010.cch.kcl.ac.uk

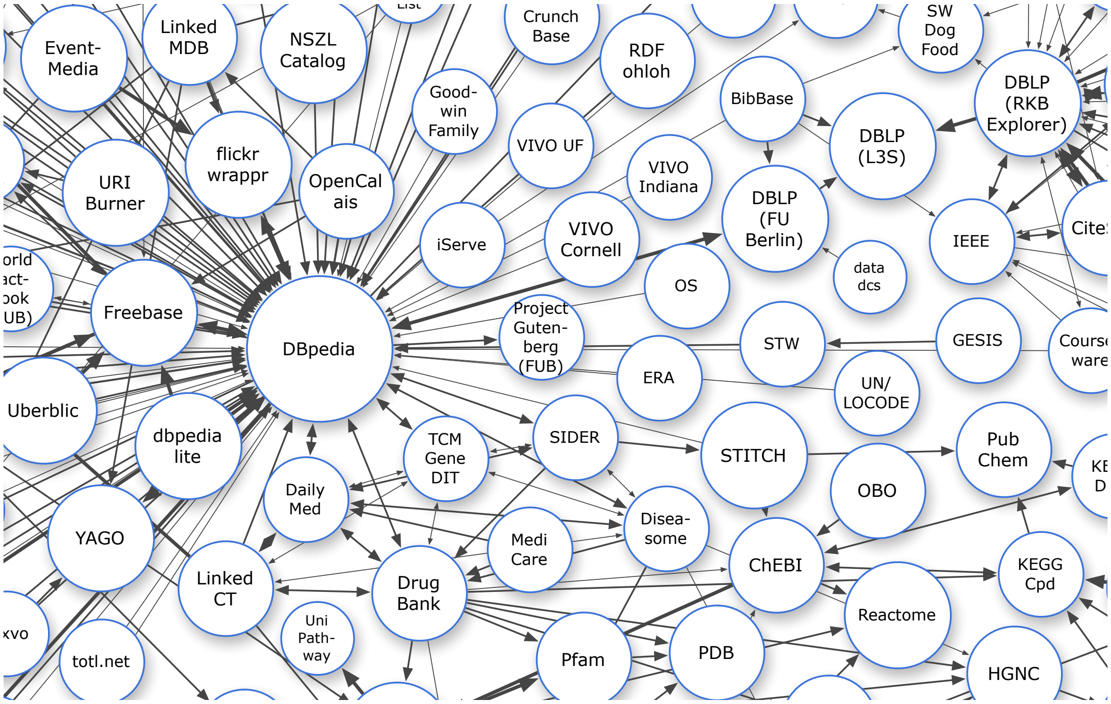

Figure 1: Small fragment of the Linking Open Data cloud diagram, by Richard Cyganiak and Anja Jentzsch as of September 2010. http://lod-cloud.net/

The proposed *LinkedHumanities* project will develop tools to (semi-)automatically integrate the highly visible Indiana Philosophy Ontology (InPhO[7]) project with some of the data repositories in the open Linked Data cloud. The InPhO project is developing techniques for building and maintaining domain ontologies. These methods are applied to the realm of digital philosophy and integrated with several partner projects. By integrating with the linked open data cloud, we hope to build a *web of facts*[19, 18] about philosophy.

Encyclopedias and other reference works have traditionally served the role of helping scholars and students to gain quick entry to the major themes of a discipline. But the encyclopedia itself has evolved in the digital age wherein some groups of scholars are creating new forms of reference work that are responsive to new research by being continually revisable. A leading project of this type is the Stanford Encyclopedia of Philosophy (SEP)[8]. The SEP is the first attempt by academics of any discipline to organize their professional subject matter by collaboratively writing, publish-

[7] http://inpho.cogs.indiana.edu

[8] http://plato.stanford.edu

ing, and maintaining a *dynamic reference work* entirely on the World Wide Web. As of November 2010, the SEP has over 1500 authors and subject editors. Those authors have submitted over 1200 entries (with an additional 250 currently on commission) containing over 14.5 million words. The authors are contributing new entries and asynchronously updating already published entries to reflect the latest developments in scholarship. Everything the authors produce (including updates) is refereed by one or more of the over 120 subject editors on the editorial board prior to publication on the web.

The approach of the InPhO project is to derive a "computational ontology"–a structured representation suitable for digital manipulation and machine reasoning–from statistical methods of information retrieval combined with expert feedback from philosophers. Recent work on these techniques has led to usable digital humanities tools and datasets, and resulted in several publications some of which were authored in a collaboration between the groups at Indiana University and the University of Mannheim [4, 14, 15, 16]. The techniques, including an innovation of our own in detecting hierarchical relationships statistically, will be improved and complemented by the data integration tools supporting the enrichment and maintenance of the ontology.

The tools we will develop in the context of the proposed *LinkedHumanities* project will allow others to employ the ever-growing InPhO dataset in their applications. For instance, a developer could use the links that connect InPhO's philosophers to their places of birth stored in Freebase to create a visual map with adjustable time-line that shows the birthplaces of important figures in philosophy over the last centuries. Heml[9], a project that builds visualization tools for historical data could benefit from the linking tools by connecting their data to other repositories. These are just some examples of how the data and the connections to other repositories could be employed in novel digital humanities applications. The principal idea of the linked data movement is to make machine-readable representations openly available and to let people build their innovative applications on top of it. We believe that the proposed project will provide the tools and datasets necessary for those projects to spring off.

As well as direct applications, we expect new research directions for Digital Humanities research to emerge from the attempt to automate the linking of machine-readable representations. Digital humanities datasets have often special properties that we will leverage for better alignment

[9] http://heml.mta.ca

algorithms. In addition, the issue of provenance is central in the digital humanities since reliable data with provenance information is necessary for most applications. Therefore, as part of the proposed project, we will investigate different data provenance schemas and expect the outcome to be highly relevant to other digital humanities projects.

History, Scope, and Duration

The current project has its origin in discussions between the Project Director Colin Allen and graduate students Cameron Buckner (Philosophy) and Mathias Niepert (Computer Science) in Fall 2005, which led to a successful application for twelve months of seed funding from Indiana University's "New Frontiers in the Arts and Humanities" program awarded in Spring 2006[10]. Work on hand-coding the ontology was begun by Buckner in Summer 2006, and Niepert started applying Information Retrieval methods to the SEP content in the Fall of that year. During Spring 2007, interfaces to gather feedback on the automated results from SEP experts and other interested individuals were prototyped, and this feedback was used to inform machine learning and machine reasoning methods to fill out the initial version of the ontology, which was then made available as a Web Ontology Language (OWL) format file.

The project subsequently received funding from the NEH in the form of a Digital Humanities Startup Grant ($29,800) during 2007-2008 and a Division of Preservation and Access award ($400,000) for 2009-2011. With this funding, the initial prototype has been developed in to a robust platform supporting individual users interested in browsing the information from the InPhO database as well as providing a number of services to other philosophy sites on the World Wide Web via an Application Programming Interface (API).[11] Technically speaking, the InPhO API capitalizes on one of the most prevalent technologies, the HTTP protocol, to enable ease of use by scholars, programmers, and scientists through nearly any web-aware interface. Each entity in the InPhO knowledge base is exposed as a resource with a unique URI which is accessed using HTTP methods, providing a consistent interface for data retrieval and manipulation. This is known as the REpresentational State Transfer (REST) paradigm of web services, pioneered by HTTP inventor Roy Fielding[9]. The InPhO data can be explored via human-friendly HTML or

[10] http://www.research.indiana.edu/recipients/winners05.html#1

[11] https://inpho.cogs.indiana.edu/docs/

in machine-friendly JSON, selected by simply adding either .html or .json to the URI of each resource. This RESTful API has been used to build an encyclopedia cross-reference engine in use by the SEP, and to access the InPhO's database of philosophy journals by the Noesis philosophy search engine[12]. The API allows efficient data transfer between different projects in digital philosophy. For instance, the Noesis project currently has no need to receive our entire ontology file when seeking to query specific pieces of information about journals. Similarly, a project wishing to trace the history of specific philosophers might only use data from our thinker database. The API allows ontology consumers to easily select only those entities and partitions of the InPhO which are relevant to their use, protecting data providers from any overexposure that may result from making large data dumps available to all-comers, while easing the processing load for data consumers who might otherwise have to parse masses of unwanted data. Individual users benefit from the modular design of the API because it allowed us to build a sophisticated interface that efficiently pulls information from several sources and combine them into a single display.

Current work is ongoing to add a bibliographic database and network visualizations to the services provided through the API. The bibliographic management system is being designed to support the needs of the SEP and InPhO, and we have initiated discussions with PhilPapers[13] about coordinating their database with ours. Our database, which already contains almost 20,000 items pulled from the SEP, will have a public interface like BibSonomy, but will support import and export of multiple bibliographic schemes, including plain text. Our approach to visualizations is being developed in collaboration with Professor Katy Börner's InfoVis and Cyberinfrastructure for Network Science groups in IU's School of Library and Information Science. In addition, the InPhO project was approached in June 2010 by Professor David Michelson of the University of Alabama about applying our methods to the digital encyclopedia on Syriac culture and history, and members of our group are now serving as consultants to that project, including deploying a version of our software on their server. This deployment for a project of direct interest to historians is necessitating that we enchance our methods for tracking data provenance.

InPhO's project director, Colin Allen, is a Professor at Indiana University with a 50% appointment in Cognitive Science (becoming Program Director in July 2011) and 50% in History & Philosophy of Science, and Adjunct Professorship of Philosophy. Allen is also the SEP's Asso-

[12] http://noesis.evansville.edu/

[13] http://philpapers.org/

ciate Editor and its Principal Programming Consultant. Mathias Niepert is a postdoctoral research associate (Akademischer Rat) in the Knowledge representation and management group at the University of Mannheim. He obtained his PhD in Computer Science from Indiana University and has extensive experience with the InPhO project. Both co-PIs bring the requisite combination of skills in computer science, especially data-integration and semantic web technologies, knowledge of philosophy, and programming experience to the InPhO project. Finally, we will exploit the twin contexts of Indiana University, with its strong School of Informatics, Cognitive Science Program, and School of Library and Information Sciences, and the KR & KM group of the University of Mannheim with its research focus on semantic web technologies.

The work proposed under this NEH proposal will last from September 1, 2011 until August 31, 2013.

Environmental Scan

Linked open data is a major research theme within the Semantic Web community. Major semantic web conferences such as ISWC and ESWC have special tracks on the topic. There is a large body of research on standards, methodologies, and applications of open linked data which we will consider and employ as part of the proposed project. One particular dataset that has contributed to the tremendous growth and interest in linked open data is the DBPedia [1] repository. DBPedia is a community effort to extract structured information from Wikipedia and to publish this information on the web. DBpedia allows users to ask structured queries against Wikipedia data, and to link other data sets on the Web to Wikipedia data. The linking of the InPhO ontology to DBPedia is one of the main objectives of the proposed project. Similar to DBPedia, we seek to create structured machine-readable data about objects and their relationships. However, our project combines information extraction algorithms with expert feedback and logic programming, and focuses on finer-grained data specific to the discipline of philosophy.

The problem of semantically linking structured data has been extensively studied in the context of database systems as duplicate detection, record linkage, and object or reference reconciliation [13]. Most existing work has focused on the design of specialized measures which estimate the similarity of objects based on their lexical properties. The Silk framework [22], for instance, combines lexical similarity measures in order to create links between objects. The use of schema information in the context of formal ontologies has only recently been proposed [20, 11]. The commonly applied methods for object reconciliation include structure-based strategies as well as strategies to compute and aggregate value similarities. Under the notion of instance matching, similarities between instance labels and datatype properties are mostly used to compute confidence values for instance correspondences. Examples of this are realized in the systems RiMOM [23] and OKKAM [21]. The participation of our object reconciliation system CODI (combinatorial optimization for data integration) has shown that our idea of leveraging schema information during the alignment process results in semantic links that are very accurate. In fact, we had the system with the highest precision values in the instance matching track.

The Heml[14] (Historical Event Markup and Linking) project is similar to ours in that it attempts to explore how disparate historical materials on the Internet can be navigated and visualized employing an XML data format. This format aims for conforming data that can be quickly parsed but provide a variety of facets on which to search for historical materials. While the *LinkedHumanities* project also aims at combining heterogeneous and disconnected collections, it leverages existing structured representations for the linking and aggregation task. The main objective is to build tools that are applicable to data from any subarea of the humanities as long as it is available in standardized linked open data formats such as the resource description framework (RDF).

Project Description, Methodology, and Standards

The *LinkedHumanities* project will be comprised of two main phases: (1) Linking digital humanities data and (2) aggregating digital humanities data. Figure 2 depicts small fragments of both the InPhO and Freebase data repositories. In both repositories, we have individual nodes corresponding to, for instance, philosophers (G.E. Moore, Ludwig Wittgenstein), documents (Principia Ethica, Tractatus), and ideas (naturalistic fallacy, private language argument). During the first phase of the project we will develop novel algorithms for linking entities in the InPhO ontology to entities in other linked data repositories. This task is non-trivial due to (a) the heterogeneity of the data and (b) the large size of the repositories which more often than not contain millions of entities. For instance, property names might vary between collections ("born on" vs. "date of birth"), individuals names might be spelled in various possible ways ("George Edward Moore" vs.

[14] http://heml.mta.ca

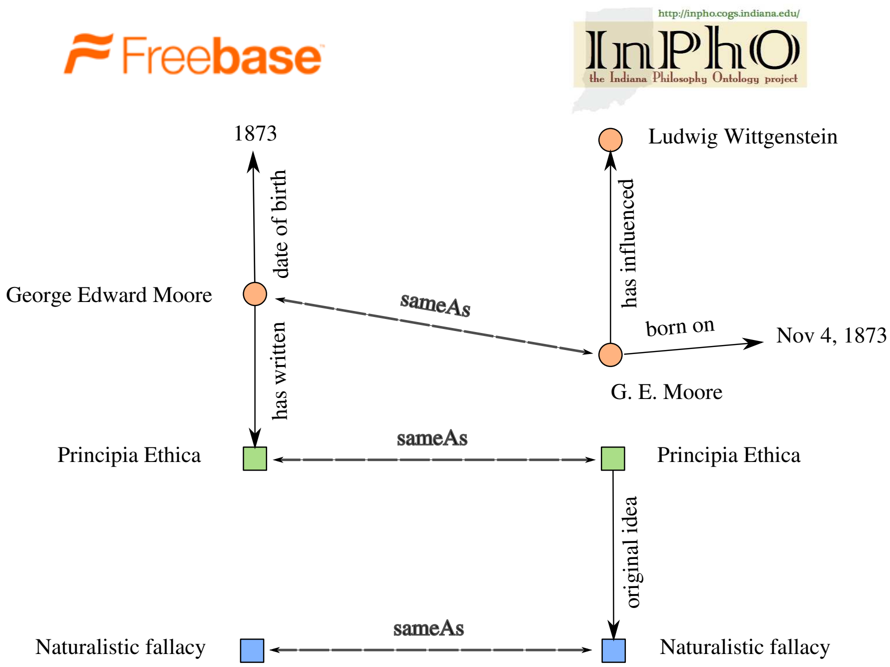

Figure 2: Linking scenario between fragments of the InPhO and Freebase data repositories.

"G. E. Moore"), and there are distinct individuals with identical names (e.g. "David Lewis"). As of November 12th, 2010, DBPedia contains 312,000 persons, 413,000 places, 140,000 organizations (including 31,000 educational institutions), and overall more than one billion facts. The enormous size of the repositories in the data cloud demonstrates that we have to design and implement highly scalable systems working with index structures and approximate algorithms such as approximate string matching [12].

Once we have implemented the alignment service that continuously creates and updates links between concepts, properties, and individuals of two knowledge bases, these links can be leveraged to enrich and populate the InPhO ontology. For instance, by creating the semantic "sameAs" links depicted in Figure 2 we can enrich the InPhO by inferring that "G.E.Moore" wrote the document "Principia Ethica." The linking of these facts which are located in different data repositories would connect the philosophers to their respective works which in turn would add additional functionality to InPhO's exploration interfaces. We also expect additional synergies and discoveries to arise from the linking of digital humanities data.

In the following we describe the methods and standards we intend to use in order to address the previously mentioned challenges.

Table 1: Current list of properties in the InPhO ontology.

Thinker Properties	Date of Birth, Date of Death, Spoke_language , Nationality, Profession
Thinker-Thinker Relations	Teacher_of, Influenced, Criticized, Defended, Dissertation_Advisor_of, Discoursed_with
Idea Properties	Idea_type [concept, position, etc.]
Thinker-Document Relations	Wrote, Edited
Thinker-Organization Relations	Member_of, Studied_at
Thinker-Idea Relations	Worked_on (problem), Created_view, Attacked_view, Espoused_view, Aware_of
Idea-Idea Relations	Opposed_to, Commits_to (idea1 commits one to idea2)
Document-Document Relations	Published_in (article in journal/book)
Document-Idea Relations	Discusses

Linking Digital Humanities Data

The integration of the InPhO ontology with the open data cloud will be the main use case of the proposed project. The InPhO is a general ontology consisting of four basic sub-ontologies: Thinker, Document, Organization, and Idea. The "Thinker" sub-ontology will classify persons of interest to philosophy along professional lines, dividing into "philosopher," "scientist," "artist," "clergy," among other categories, and was bootstrapped from information found in the SEP itself as well as public databases such as the Philosophy Family Tree[15]. The "Document" and "Organizations" sub-ontologies are largely taken (with appropriate augmentation and pruning) from standard ontology libraries (such as the Protégé or AKT ontology libraries). The current version of the ontology also contains several object and data properties. A list of these properties is provided in Table 1.

While linked open data avoids many problems connected with the use of expressive ontologies such as the knowledge acquisition bottleneck, data heterogeneity remains a challenging problem. In particular, identical objects may be referred to by different URIs in different data sets. The goals of the first phase of the proposed project are the linking of InPhO's non-taxonomic relations to the equivalent relations in the repositories we want to integrate our data with. For instance, we have to identify the relations in DBPedia and Freebase that are equivalent to InPhO's "has influenced" relation. Since there is only a small and limited number of such relations we plan to accomplish

[15] https://webspace.utexas.edu/deverj/personal/philtree/philtree.html

this task semi-automatically. We will develop algorithms that suggest a set of candidate relations which will then be validated by human experts.

Once we have an alignment of the properties we can begin the actual linking process between two data repositories A and B. Due to the large scale of the datasets we first have to extract a set of "alignment candidates." We will do this using existing approximate string matching and indexing methods from the area of information retrieval. One particular candidate we will evaluate in the context of our project is the FLAMINGO package[16][2]. The objective is to extract, for each entity in dataset A, a reasonable number of candidates in dataset B and vice versa. We envision a number between 1 and 100 candidates per individual but this number is influenced by the complexity of the alignment algorithms that we will apply in later phases and the size of the respective datasets. The structured nature of the data also allows us to use additional methods for generating alignment candidates. For instance, assume we have an object in dataset A whose "name" label is missing. Of course, this would render the method of approximate string matching on names useless for the generation of candidates. Nevertheless, since we have additional properties of the object (e.g., date of birth) we can find all individuals in dataset B that have the same or similar values for a selected subset of the properties. This way the data integration system will be able to find links between individuals in the presence of uncertain and incomplete data.

With the collected set of alignment candidates at hand the actual linking algorithm can be applied. Its output will be a set of links that *reconcile* identical objects contained in both A and B. Contrary to the candidate generation process the emphasis here is on high precision results. The problem of object reconciliation has been a topic of research for more than 50 years. It is also known as the problem of record linkage [7], entity resolution [3], and instance matching [8]. While the majority of the existing methods were developed for the task of matching database records, modern approaches focus mainly on graph-based data representations extended by additional schema information. The University of Mannheim's KR & KM group is actively working on algorithms that employ both logical reasoning and linear optimization techniques to take into account the *overlap* of derivable properties of objects. To demonstrate the concept of overlap, let us consider the example in Figure 2 again. Most string similarity measures would deem the strings "George Edward Moore" and "G.E. Moore" only moderately similar. However, since we also know

[16] http://flamingo.ics.uci.edu/releases/4.0/

that the year of the "date of birth/born on" properties is the same for both objects, the likelihood of linking these two objects increases. This overlap information is combined with the classical similarity-based approach in a novel framework for object reconciliation [17]. As part of the proposed project, we will adjust and extend the existing object reconciliation system CODI for the task of linking digital humanities data. CODI's participation in the most recent OAEI[17] has shown that the system produces instance alignments with higher precision than any of the other existing systems and that it generates high quality alignments between datasets like Freebase and DBPedia.

The final stage of this phase will consist of the employment of the developed linking system as a continuously running data integration service. This service will create novel links and correct existing ones. In addition, the algorithms will be evaluated on digital humanities datatsets other than InPhO. Furthermore, we will also design user interfaces and integrate those with the InPhO community system. Currently, InPhO community members can help to extend and populate the ontology by providing feedback facts about philosophical ideas and properties of philosophers. We will complement the existing interfaces with those that allow InPhO's users to verify or falsify link recommendations derived by the integration system. The InPhO project staff has experience with the design of user interfaces and, therefore, the modifications to the InPhO community software will be rapidly executed.

The generated semantic links will be made publicly available in a RDF repository with SPARQL[18] endpoint. In addition to the link repository, we will also make available provenance information (e.g., whether the user who verified the link was an expert or interested amateur, two of InPhO's user categories) and a confidence value computed by the data integration algorithm indicating the degree of correctness of the link. This information will be useful for other projects that wish to build their own applications on top of the provided link data.

Aggregating Digital Humanities Data

The second phase of the project will harness the previously created links to populate the InPhO ontology. To demonstrate how this would work, consider the "has written" relationship that holds between the objects "George Edward Moore" and "Principia Ethica" in Figure 2. Once the depicted "sameAs" links have been generated between the knowledge bases, the integration system

[17] http://oaei.ontologymatching.org/2010/

[18] http://www.w3.org/TR/rdf-sparql-query/

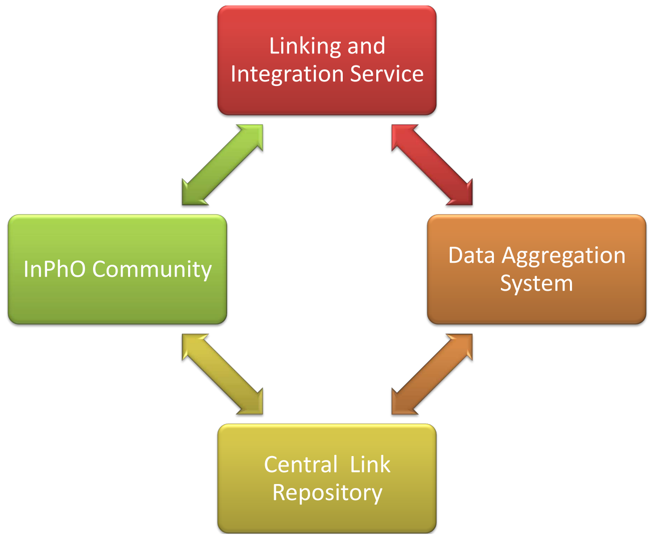

Figure 3: Flowchart of the proposed linking and integration system.

can immediately infer and add the "has written" relationship between "G.E. Moore" and "Principia Ethica" to the InPhO ontology. Moreover, it can also infer that "George Edward Moore" introduced the idea of "Naturalisitc fallacy" which would not have followed from either repository in isolation. This simple example shows how the alignment of objects in different datasets can lead to a tremendous amount of novel knowledge.

Currently, the InPhO ontology contains some instances for the non-taxonomic relations such as "date of birth" or "has influenced." Based on the generated links from the data integration service we will build algorithms that continuously add novel instances and properties to the InPhO ontology. Data provenance will be an important component of the methods developed in this phase. Due to the InPhO ontology's high quality standard we need to store additional information about the origin of the inferred properties and instances in the ontology. We are currently investigating different existing approaches for this task. Fortunately, the KR & KM group also has expertise in the research area of provenance for metadata [6, 5] and some of its members are active in the W3C Provenance Incubator group[19]. We envision the publication of the provenance data as part of the RDF repository alongside the primary data.

Figure 3 depicts the flow of information between the different modules of the proposed system. The linking and integration service loads the current version of the InPhO ontology in order to

[19] http://www.w3.org/2005/Incubator/prov/charter

create and correct semantic links to other repositories in the open data cloud. The generated links are in turn presented to the members of the InPhO community for verification. The individual semantic links are stored along with their provenance data in the central link repository. Based on the semantic links generated by the linking and integration service and the data in the cental link repository, the data aggregation component infers novel facts and adds those to the repository and the InPhO ontology. The latter step does not require user interaction since the provenance data for newly added facts can be derived from the provenance data associated with the semantic links. For instance, assume that we know the confidence values and IDs of users who have verified the "sameAs" links depicted in Figure 2. Then we can annotate inferred facts with the provenance data of the links that were used in the derivation. There exist a general theory of provenance computations [10] that we will investigate in the context of the proposed project.

The only risks the proposers see at the moment are those stemming from (a) scalability and (b) data heterogeneity challenges. Both potential problems have been addressed in the previous description of the methods. Both Co-PIs have experience in working with digital humanities data and know about their specific properties such as heterogeneity and incompleteness. Co-PI Allen is the director of the InPhO project and has been overseeing its development and research efforts over the past years. In addition, the KR & KM research group is a leader in the design and development of semantic web technologies including the development of distributed reasoning algorithms. Co-PI Niepert has been working for many years on algorithms that have to strike a balance between scalability and sophistication. In summary, we believe to be well positioned to meet the potential challenges of the proposal and to minimize the risks of the project.

Organization of and Access to Material

Most of the data of the InPhO project are stored in relational databases which can be queried through a RESTful API[20]. Additionally, we also provide monthly OWL dumps of the ontology. As part of the proposed *LinkedHumanities* project, we will store the instances of the InPhO ontology and the generated semantic links in an RDF repository. This RDF repository will be accessible through a SPARQL endpoint, an API (providing query results in the JSON format), and the ontology itself will also be published in OWL 2. The source code and documentation of the data

[20] https://inpho.cogs.indiana.edu/docs/

integration and aggregation tools will be made publicly available on the project webpage.

Intellectual Property Rights and Copyright

To the best of our knowledge, all repositories in the linked open data cloud (for a fragment of the LOD cloud see Figure 1) are published under a creative commons license meaning that others are free to copy, distribute, adapt, and transmit the data as long as they attribute the creator and/or distribute the tranformed data under the same or a similar license. The two repositories we have mentioned, Freebase and DBPedia[1], are both available under licenses CC-BY[21] and CC-BY-SA[22], respectively. In fact, the creators of the datasets explicitly encourage the reuse and distribution of the data. The same is, of course, true for the InPhO data and software. All software we will use is either developed during the course of the proposed project, was developed as part of the InPhO project, or is otherwise in the public domain.

Work Plan

A comprehensive work plan indicating benchmarks and milestones is included as an attachment.

Nature of Collaboration and Staff

Co-PI Colin Allen is director of the InPhO project and oversees the development of the project's infrastructure and algorithms design. Co-PI Mathias Niepert is also co-founder of the InPhO project and is an expert on probabilistic-logical approaches to data integration. The collaboration between these two groups would continue to combine the KR & KM group's strengths in data integration and Semantic Web research with IU's strong expertise in philosophy and digital humanity projects.

The Co-PIs have a proven record of collaboration within and between institutions, leading to technological developments and published papers and conference presentations. The InPhO project was jointly conceived and executed while Niepert was at Indiana, and since moving to the Mannheim, he has independently developed new lines of research based on the original ideas (e.g., [4]). Allen's selection in 2010 for a Research Award from the Alexander von Humboldt Founda-

[21] http://creativecommons.org/licenses/by/2.0/legalcode

[22] http://creativecommons.org/licenses/by-sa/3.0/legalcode

tion[23] ensures his regular presence and academic connections with researchers in Germany. Although, the cross-institutional collaboration has not yet involved any shared funding, there is every reason to think that the terms of the NEH/DFG awards will minimize administrative problems, enabling us to greatly enhance our already successful collaboration, and allowing us to continue to develop new directions of research in computer science and digital humanities. Both Allen and Niepert have also established themselves as effective advisors for students working on these projects, and the exchange visits between the two sites will enhance collaboration at all levels. It is worth mentioning, also, that IU undergraduate Jaimie Murdock, who is responsible for design and implementation of the InPhO API has applied for Fulbright funding to study in Germany in Fall 2011, holding the further prospect of strengthening this German-U.S. alliance.

Dissemination

The results of this project will be disseminated in three ways:

1. All software, data, and its products will be made freely accessible through (a) open source of the code, where appropriate, (b) dynamic accessibility to data via an open-access API, and (c) the provision of standards-compliant data structures such as OWL and RDF to support the Linked Data initiative. We will also work with developers of 3rd-party academic projects such as Noesis and PhilPapers, to ensure that their data needs are being met by our systems.

2. Methods and results will be published in the scientific and philosophical literatures, through the proceedings of prestigious conferences in computer science such as JCDL, ISWC, ESWC, and KEOD, and in high-visibility philosophy journals such as *Synthese*.

3. We will continue to construct and enhance web portals that enhance scholarly and public access to the humanities. We will do this directly for philosophy through the InPhO website, and by links from the SEP and Noesis. (See the appendix for screenshots of a working demo of new links proposed for all SEP articles to entry-specific pages at InPhO and PhilPapers.) We will also do this indirectly for other humanities subjects by making our expertise available to other projects in the digital humanities, such as the aforementioned digital encyclopedia of Syriac history and culture.

[23] http://www.humboldt-foundation.de/pls/web/pub_laudatio.main?p_lang=de&p_fgb=1B