8-28-2012

Facilitating Knowledge Sharing and Analysis in EnergyInformatics with the Ontology for Energy Investigations (OEI)

Michael Cotterell
University of Georgia, mepcott@uga.edu

Jie Zheng
University of Pennsylvania, jiezheng@pcbi.upenn.edu

Qin Sun
Georgia Institute of Technology, qinsun@gatech.edu

Zhengkai Wu
Georgia Institute of Technology, zkwu@gatech.edu

芽|Sprouts Working Papers on Information Systems ISSN 1535-6078

Facilitating Knowledge Sharing and Analysis in Energy Informatics with the Ontology for Energy Investigations (OEI)

Michael Cotterell
University of Georgia, USA

Jie Zheng
University of Pennsylvania, USA

Qin Sun
Georgia Institute of Technology, USA

Zhengkai Wu
Georgia Institute of Technology, USA

Craig Champlin
Colorado School of Mines, USA

Aaron Beach
National Renewable Energy Laboratory, USA

Abstract
Just as the other informatics-related domains (e.g., Bioinformatics) have discovered in recent years, the ever-growing domain of Energy Informatics (EI) can benefit from the use of ontologies, formalized, domain-specific taxonomies or vocabularies that are shared by a community of users. In this paper, an overview of the Ontology for Energy Investigations (OEI), an ontology that extends a subset of the well-conceived and heavily-researched Ontology for Biomedical Investigations (OBI), is provided as well as a motivating example demonstrating how the use of a formal ontology for the EI domain can facilitate correct and consistent knowledge sharing and the multi-level analysis of its data and scientific investigations.

Keywords: Energy Informatics, Ontology design, Knowledge modeling

Permanent URL: http://sprouts.aisnet.org/12-4

Reference: Cotterell, M., Zheng, J., Sun, Q., Wu, Z., Champlin, C. , Beach, A. (2012). "Facilitating Knowledge Sharing and Analysis in Energy Informatics with the Ontology for Energy Investigations (OEI)," Proceedings > Proceedings of Energy Informatics . *Sprouts: Working Papers on Information Systems*, 12(4). http://sprouts.aisnet.org/12-4

INTRODUCTION

Shared semantic vocabularies and ontologies allow separate information systems to integrate at the conceptual level, enabling analysis tools and programs to infer the meaning of data without it being hard-coded. Formal ontological structures represent high-level semantic concepts and their relationships in a standard (and formal) way that can be shared and used by different domain-specific ontologies. For example, a top level ontology might represent the concepts and relationships associated with the temporal and spatial properties of things. This high level ontology could then be shared by domain-specific ontologies thus enabling the inter-domain sharing of analysis tools and programs that utilize the concepts and relationships from this high level ontology. In order to share a high-level ontology with other domains a domain-specific ontology or vocabulary must root its semantic concepts in those from the high-level ontology. This rooting is sometimes called "alignment." Extending the example previously mentioned, this would mean that the spatial and temporal concepts in the domain-specific ontology are rooted/related to those high-level concepts in the shared formal ontology.

The benefits of such structuring of ontologies has been demonstrated in the biomedical community in which the Ontology for Bioinformtics Investigations (OBI) is rooted in the Basic Formal Ontology (BFO) (Brinkman et al., 2010; Pisanelli, 2004). While many fields related to energy systems have adopted semantic technologies to develop shared vocabularies and ontologies, these have not been rooted in a basic ontology that can be shared with other domains. This paper presents an ontology designed to provide a core structure that helps relate domain-specific ontologies and vocabularies dealing with energy systems back to the BFO, taking advantage of ontological structures already developed in the biomedical community that also apply to energy systems.

This paper describes the development of the Ontology for Energy Informatics (OEI) with the goal of demonstrating the utility of such of rooting energy-system related ontologies in a shared ontology. In this paper, we discuss examples of how the National Renewable Energy Laboratory (NREL) is integrating this approach in its own data processing systems to encourage generative development and sharing of its energy models and software. OEI is at the core of NREL's Energy Systems Integration (ESI) efforts to integrate electricity, thermal, fuel and data systems at all scales across the grid.

The rest of this paper is organized as follows: the remainder of this section will describe, in general, ontologies as well as provide some background information on the foundations for OEI; the next section provides an overview of OEI; then we include some motivating examples for ontologies in the EI domain; after that, we provides an example of a real-world application demonstrating the value of ontologies in the EI domain with respect to knowledge sharing and multi-level analysis; in the next section, we present some related work; and, conclusions and future work are described in final sections.

Note. Throughout this paper, terms referring to ontological concepts will be written in italics (e.g., *photovoltaic conversion device*, which refers to the *device* that is more commonly known as a solar cell).

Energy Informatics

The idea of Energy Informatics (EI) is to increase the efficiency of energy demand and supply systems through scientific investigations (Watson and Boudreau, 2011). In particular, the overall domain of EI concerns itself with the appropriate analysis of of available information in order to optimize the performance of these systems. According to Watson and Boudreau (2011), EI can be represented by the equation:

$$\text{"Energy} + \text{Information} < \text{Energy"}$$

The idea is that we can make better decisions about how to both use and conserve energy through the use of information.

Just as with other Informatics domains, EI has its own investigations. These scientific investigations involve their own basic entities, protocols, instruments, materials, type of data generated and consumed, and the types of analyses performed, etc (Smith et al., 2007). Information related to these concepts make up some of the key components of information systems designed to facilitate Energy Informatics. In this paper, we will describe how these concepts can be organized and formalized using an ontology as well as how such an ontology can facilitate multiple dimensions of analysis in information systems.

Ontologies

“Ontologies” are formalized, domain-specific taxonomies or vocabularies that are shared by a community of users (W3C OWL Working Group, 2009b). They help facilitate interoperability among domain-specific information systems as well as promote reuse of domain knowledge by providing a common vocabulary and set of relationships for a particular domain (Pinto and Martins, 2004). Without ontologies, domain experts often find themselves using different terms to explain the same concepts. Ontologies help resolve such idiosyncrasies.

In formal ontologies, concepts are not only specified by giving them definitions or descriptions but by also describing their contexts or how they relate to other concepts (Spear, 2006). It is in this way that ontologies differ from terminologies (i.e., simple lists of terms with their associated definitions). This is beneficial because it helps facilitate the sharing of knowledge by ensuring that everyone is referring to the same set of concepts and their relationships. For example, when an ”instance”–a piece of information that realizes a concept in an ontology–is tied to a particular concept in an ontology (e.g., an instance of a solar panel corresponding to a *photovoltaic conversion device*), a domain expert can use the vocabulary and relationships defined in the ontology to understand the context of the information with respect to the domain. Also, if two people are trying to convey the same piece of information with different terms, mapping that information to a concept in an ontology will help disambiguate the meaning behind the information (Spear, 2006). This is particularly useful when information is communicated between computer systems in a machine-processable and machine-interpretable way as ontologies ensure that semantic “annotations”–mappings to ontological concepts–are consistent to both the annotators, or domain experts, and the systems that take advantage of those annotations (Bussler et al., 2002). Additionally, information systems can use logic-based context reasoning to check the ontology and related instances for consistency as well as infer higher level contexts from the information (Wang et al., 2004).

With respect to the Semantic Web, in order to facilitate accessibility and use of semantics, ontologies are often implemented in the W3C-endorsed Web Ontology Language (OWL) 2 (Grau et al., 2008; W3C OWL Working Group, 2009b). A common tool for creating OWL ontologies is Protg (Gennari et al., 2003), an open source ontology editor developed at the University of Stanford's Center for Biomedical Informatics Research. OWL 2 is an extension and revision of the 2004 OWL 1 specification, and is designed to facilitate ontology development and sharing via the Web and the Resource Description Framework (RDF). RDF is a standard model for data interchange on the Web (W3C, 2009). A standard called OWL 2 DL (W3C OWL Working Group, 2009a) exists that helps maximize the expressiveness of an OWL ontology without losing computational completeness or decidability when combined with reasoning systems (i.e., facilitating what is kown as direct semantics). According to the correspondence theorem in (W3C OWL Working Group, 2009c), there is a well documented and precise relationship between the direct semantics that are possible with OWL 2 DL and the semantics that are possible when the instances are mapped onto an RDF-based graph for the Semantic Web. Another advantage of facilitating the reasoning capabilities though the use of OWL 2 DL (e.g., with Pellet (Sirin et al., 2007) or HermiT (Shearer et al., 2008)) is that an ontology engineer can expect to maintain a proper level of ontological consistency and reliability (Bock et al., 2008).

The goal of this project is the creation, maintenance and application of an ontology for the EI domain and its investigations, OEI, that is both complete (i.e., minimize term ambiguity) and valid (i.e., every ontological concept is connected to every other term through one or more object property relationships) in order to facilitate the informatics-related tasks of knowledge sharing and multi-level statistical analysis.

Foundations and the Ontology for Biomedical Investigations (OBI)

The foundation of OEI is an extension of existing ontologies. Providing the basis for general scientific investigations is the Ontology for Biomedical Investigations (OBI). OEI is built by extending a semantically valid and complete subset or view of OBI that is composed of concepts that are relevant to general scientific investigations. The following section provides some background on OBI as well as some of reasons for why it was chosen to form the foundation for OEI.

OBI is being developed to support the consistent annotation of biological and clinical investigations regardless of the particular field of study (Brinkman et al., 2010). It covers concepts that describe all aspects of an investigation, including materials, protocols, assay, devices used in assay, generated data and types of analysis applied to the data. OBI is based on the Basic Formal Ontology (BFO) (Arp and Smith, 2008; Pisanelli, 2004) and follows Open Biomedical Ontologies (OBO) Foundry principles. It is inter-operable with other biomedical ontologies under the OBO Foundry umbrella since they have shared common foundation ontology, BFO, and use a common set of relations. OBI uses the Information Artifact Ontology (IAO)[1] as a middle tier ontology to describe information, such as data, document, and design. IAO was created by OBI developers to represent information concepts out of OBI scope and generally applicable for all knowledge domains.

Although OBI aims to describe biomedical investigations, it contains generic concepts that can semantically represent every scientific investigation in all domains (see Figure 1). It also covers

[1]IAO: `http://code.google.com/p/information-artifact-ontology/`

general statistical analysis (e.g., error correction, data transformation and statistical hypothesis tests) as well as many instruments that are also used in energy investigations.

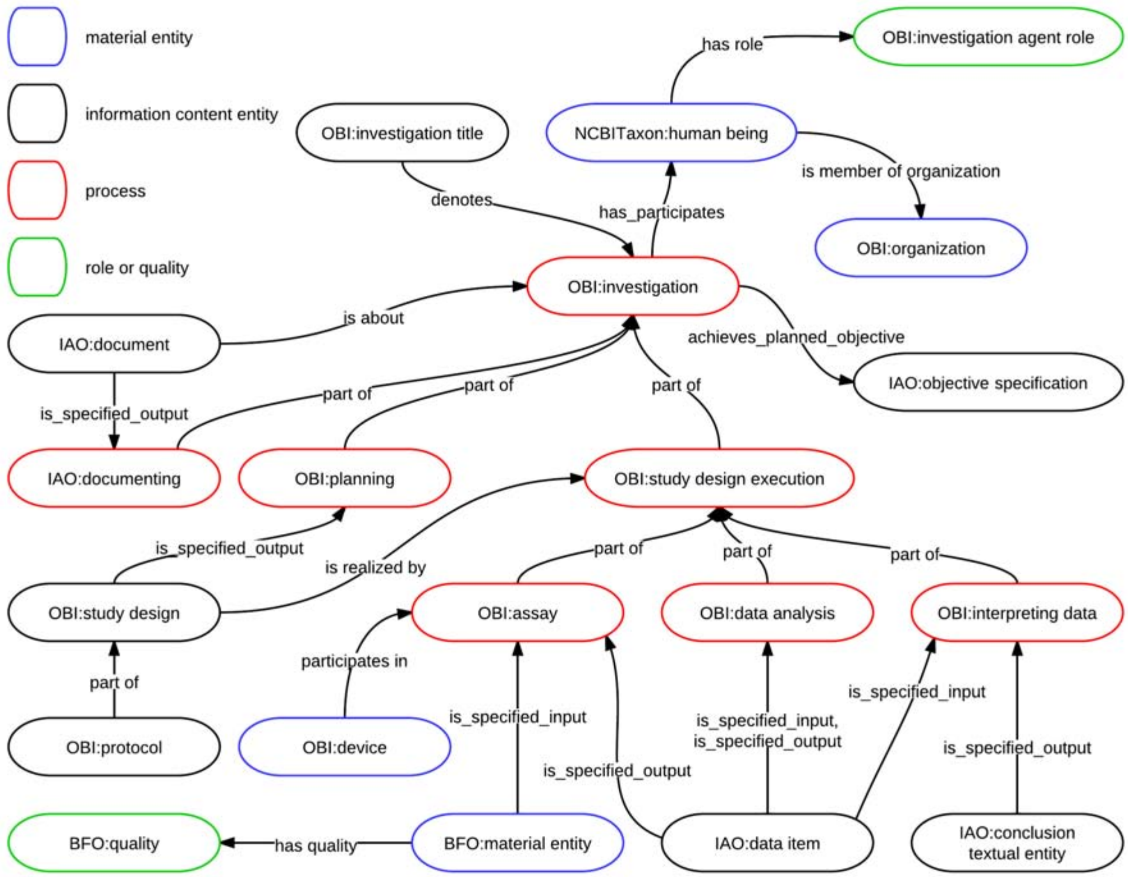

Figure 1: Semantic Representation of Scientific Investigation Using OBI

In OBI, an *investigation* has a *title*, a specific *objective* and is performed by a *human being* who is a member of some *organization* and has an *investigation agent role* (as seen in the upper left part of Figure 1). It contains three main processes, *documenting*, *planning* and *study design execution* (as seen in the middle part of Figure 1). The *Documenting* process generates a *document* about the *investigation* (as seen in Figure 1), *planning* how to perform the investigation and generating a *protocol* that can be followed in the *study design execution* (as seen in the center of Figure 1), and a *study design execution* that executes the *study design* by performing *assay* on some materials using specific instruments (*device*), generating data about the material, the *data analysis* of generated data in the assay, and the conclusions drawn from the investigation (*interpreting data*).

Due to the amount of concept overlap, building OEI on top of OBI will help eliminate redundant work and speed up the development of the ontology. Besides, OEI will inevitably deal with certain biology contexts (e.g., biomass generation), and will, therefore, benefit from many of the existing biological (and perhaps biomedical) relationships inherited from OBI. Extending OBI and building on the basis of BFO will allow OEI to reuse generic concepts defined in the OBO foundry ontologies, such as environmental material related concepts defined in EnVironment

Ontology (EnVO)[2] and metrical units defined in Units of Measurement (UO)[3].

OEI OVERVIEW

Although OEI is still early in its inception, the overall development process is being structured according to the ontology development life-cycle model mentioned in (Pinto and Martins, 2004). In order to promote ontology reuse, the "composition/integration" approach was chosen for composing OEI. In this approach, in addition to adding new terms that convey knowledge in the specified domain, other ontologies are extended and aggregated in order to promote reuse and formalize cross-domain alignment (Pinto et al., 1999). Specifically, as mentioned in Section , OEI is built by extending a subset of the concepts and relationships contained within OBI and other formal ontologies.

Knowledge Extraction & Conversion

In order to create the structure for the ontology in such a way that knowledge extraction is facilitated, energy-related concepts and relationships have been extracted from existing semantic sources available in the EI domain. Specifically, the following two semantic wikis have served as excellent sources for OEI terms:

- *OpenEI*: The Open Energy Information initiative, found at `http://openei.org/`, refers to itself as a free, open source knowledge-sharing platform for data, models, tools, and information related to clean and renewable energy systems (Brodt-Giles, 2012; Young et al., 2012). It's built on top of the Semantic MediaWiki platform which enables its information to be exposed as an RDF graph for semantic information interchange purposes such as Linked Open Data (Krötzsch et al., 2006). It is sponsored by the U.S. Department of Energy and developed at the National Renewable Energy Lab (NREL).

- *Reegle*: Reegle, found at `http://www.reegle.info/`, describes itself as a portal for clean energy information that targets specific stakeholders, including governments, project developers, businesses, financiers, NGOs, academia, international organizations and civil society (Shi, 2010). Of particular interest is Reegle's extensive glossary which links all of its terms according to their relationships (Bauer et al., 2011).

Existing ontologies have already been created for the energy domain and would also serve as an excellent source for EI-related concepts, however, these ontologies are often not available for download, lack the foundation of a formal high level ontology, or both.

In order to speed up and facilitate collaborative development of OEI, the Protégé (Zhao et al., 2012) program is being used. It allows us to take advantage of the full OWL 2 DL specification in an easy-to-use way.

---

[2]EnVO: `http://environmentontology.org/`
[3]UO: `http://code.google.com/p/unit-ontology/`

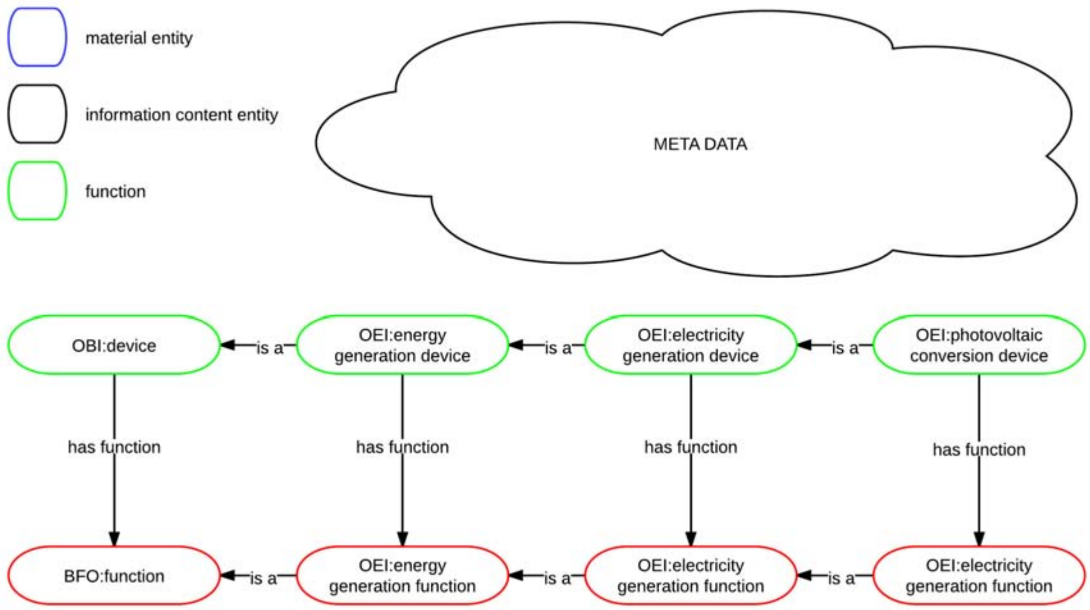

Figure 2: Extended OBI *device* concept in OEI

Concept Hierarchy & Object Properties

The concepts in OEI are structured and organized in such a way that they maintain alignment with the existing ontological terms in the subset or view of OBI that serves as the ontology's formal foundation. The descriptions for two of the OEI-specific concepts, *photovoltaic conversion device* and *smart electricity meter*, in Manchester OWL Syntax (Horridge et al., 2006) can be seen in Figures 3 & 4, respectively. These two terms are good examples because a different approach was used to define each. For *photovoltaic conversion device*, an equivalence relation was defined. This says that anything is both a *device* and *has function* some *photovoltaic conversion function* is also a *photovoltaic conversion device* and a subclass of *of energy generation device*. For *smart electricity meter*, no such equivalence class was defined. In the future, when the characteristics and relationships of these terms are better understood within the contexts of each other, the asserted relationships can be refined. For now, we have chosen to model more general concepts using equivalence relations and more specific concepts using direct sub-classing. Additionally, some of the object property relationships for *photovoltaic conversion device* can be see in Figure 2.

Motivating Examples

In the following subsections, we explore some of the motivating examples that demonstrate how the EI domain can benefit from the realization of a formalized domain ontology.

Linked Open Data

Linked Open Data (LOD) refers to the way in which semantically structured data is published and connected on the Semantic Web (Heath and Bizer, 2011). The goal is to harmonize data from

Class: *photovoltaic conversion device*

EquivalentTo:
device and (*has function* some *photovoltaic conversion function*)

SubClassOf:
energy generation device

SubClassOf (Anonymous Ancestor):
has function some *electricity generation function*
has function some *energy generation function*
has function some *function*
is specified output of some *material processing*
material entity and (*is specified output of* some *material processing*)

Figure 3: Concept description for *photovoltaic conversion device* in Manchester OWL Syntax.

Class: *smart electricity meter*

SubClassOf:
electricity meter

SubClassOf (Anonymous Ancestor):
has function some *electricity measure function*
has function some *function*
is specified output of some *material processing*
device and (*has function* some *measure function*)
material entity and (*is specified output of* some *material processing*)

Figure 4: Concept description for *smart electricity meter* in Manchester OWL Syntax.

different resources and allow users of the data to issue logic-based queries (e.g., using the SPARQL query language (W3C, 2008)).

Within government, considerable interest in LOD has grown in recent years. The United States' Open Government Initiative has created a site, found at `http://semantic.data.gov/` to serve as a LOD hub for the open and semantic data made available through various channels of the government (Hendler et al., 2012). Interest has even fostered within the EI domain itself. The Energy Data Collection (EDC) Poject, funded through the Digital Government Research Center (DGRC), seeks to expose energy data using LOD and the Semantic Web (Ambite et al., 2000; Hovy, 2003).

Suggestions in Service Composition

Service composition–combining multiple services into some kind of workflow–is a task that becomes more and more difficult as the complexity of the workflow increases and as the number of inputs and outputs for each service increases. This process is also difficult due to the lack of semantics surrounding the data involved. For example, just because an output to one service and the input of another service are both labeled as a temperature with some numeric type constraint, that doesn't necessarily mean they represent the same lower level concept (McArthur et al., 2007). It is entirely possible that one could represent temperature in Celsius and the other in Fahrenheit. Imagine if the the inputs and outputs were complex XML structures. Then the difficulty of matching up outputs to inputs during the composition process increases even more.

However, this whole process, as demonstrated thoroughly in the biomedical domain, becomes easier when performed in a semi-automatic fashion using semantics (Wang et al., 2011). By annotating the inputs, outputs, and function of each service with concepts from an ontology, you make it possible for algorithms to help suggest where services should be placed in the workflow (Dhamanaskar et al., 2012). These algorithms take advantage of the concept hierarchies and shared vocabulary contained the ontologies do determine both what inputs and outputs should be matched (sometimes referred to as the impedance mismatch problem) and in what order the services should be composed in order to achieve the goal of the workflow. One of the goals of OEI is to serve as the ontology that facilitates these benefits for services that exist within the EI domain.

Ontology-Driven Model Development

In general, how a simulation model is designed is dependent on the objective or purpose of the simulation. According to Benjamin et al. (2006), the requirements elicitation process for a simulation model can be benefited through the use of ontologies. The shared vocabulary and conceptual relationships in an ontology make it possible for one to specify some of the minimal model requirements and infer what else is necessary for the overall analysis.

An ontology for discrete-event modeling and simulation called the Discrete-event Modeling Ontology (DeMO) already exists with this goal in mind (Silver et al., 2011). By coupling domain-specific ontologies with concepts from DeMo, the process of creating domain-specific models for simulation purposes becomes easier (Miller et al., 2004). OEI can be used with DeMO and other similar ontologies to facilitate the same ontology-driven model development within the EI domain.

REAL-WORLD APPLICATIONS

For the purposes of this paper we will explore applicability of OEI when combined with the Semantic Data Integrator (SDI), a generic framework for collecting, describing and processing time-series data using Web services and semantic annotations. SDI is developed at the National Renewable Energy Laboratory (NREL), and it is designed to make the meter data for the entire NREL campus machine-readable via the Web.

Semantic Web Services & SDI

Web Services can be described using semantic annotations that enable other Web services or applications to integrate their capabilities. For instance, a meter that measures the power generated by a photovoltaic (PV) system might send its raw data to the SDI system. This raw data stream might be annotated as coming from a meter at a certain location via a particular communication protocol. A Web service might then be made which cleans error values caused by the data collection protocol or the meter itself. This Web service can be described in terms of what it does (cleans error values) and to what types of data streams it applies (raw data streams collected for a particular communication protocol or meter model). Another program may use historical PV generation, along with weather forecasts, to predict future PV generation profiles. This program can specify that it requires the historical PV generation at a certain location, cleaned of errors values. Using the semantic annotations, SDI identifies which raw data streams and data processes are required to achieve the requested data (cleaned historical PV generation data). Figure 5 depicts the different aspects of this process, showing how the annotation of the raw data stream along with a data cleaning Web service allows a PV forecasting program to make its request using a semantic query without knowing the exact ID of the data it needs.

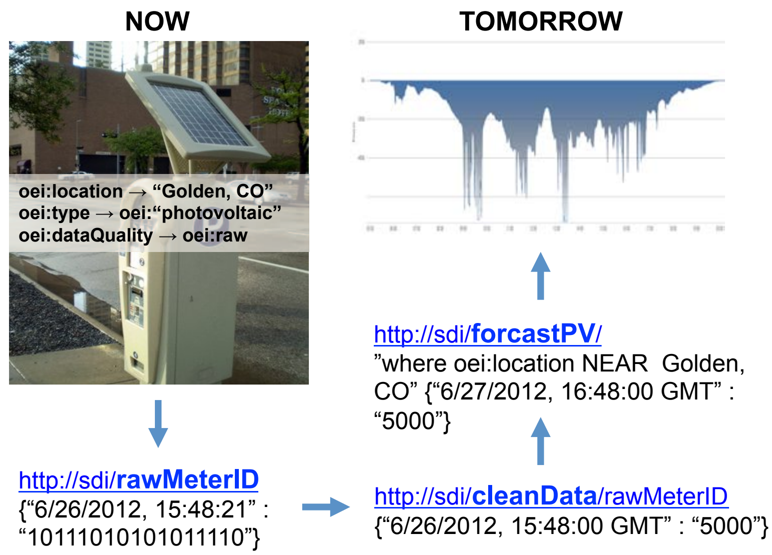

Figure 5: Example of PV data being collected, cleaned and used for foracasting

The ability to semantically describe the input, output and operation of Web service is enabled via the Web Service Description Language (WSDL) (Chinnici et al., 2009) and Web Application Description Language (WADL) (Hadley, 2009). Using these languages the function of a Web service along with how it might be connected with other Web services and data sources may be inferred from the semantic annotations (McIlraith et al., 2001). Combining energy models and analysis tools with Web Services and OEI enables energy researches to describe their own systems in a way that is translatable to other researchers and domains without making assumptions about how the data will be shared or used by other systems.

Knowledge Sharing

There exists a successful bridge between knowledge sharing and consistent references to terms in an ontology by way of the semantic annotations included in semantic Web services. For example, the output of Web services with such annotations can directly be converted to RDF triples for exposure in the Semantic Web and Linked Open Data. SAWSDL, an extension to WSDL (Farrell and Lausen, 2009; Kopeckỳ et al., 2007; Verma and Sheth, 2007). An example of this is the annotated time-series output of SDI. Each piece of information is annotated with a concept from OEI. These concepts have relationships to other ontological concepts within OEI and other ontologies. Many of these other ontologies expose their instance data in RDF triple forms for exposure on the Semantic Web.

Facilitating Multi-level Analysis

Successful annotation of data using a formal ontology such as OEI enables multi-level analysis. An example of such an analysis in the biomedical domain can be seen in Hu et al. (2009) and Zheng et al. (2011). In this section we will explain how the EI domain can also benefit from this kind of functionality.

Take, for example, the concepts in OEI that are related to time series data. From a hierarchical point-of-view, all such ontological terms are organized in much the same way (i.e., there conceptual object relationships are structured in a uniform fashion) and, therefore, can potentially be analyzed the same way. The terms *time sampled electricity measurement data set* and *time sampled thermal measurement data set* from OEI, as seen in Figure 6. Although this may seem like comparing

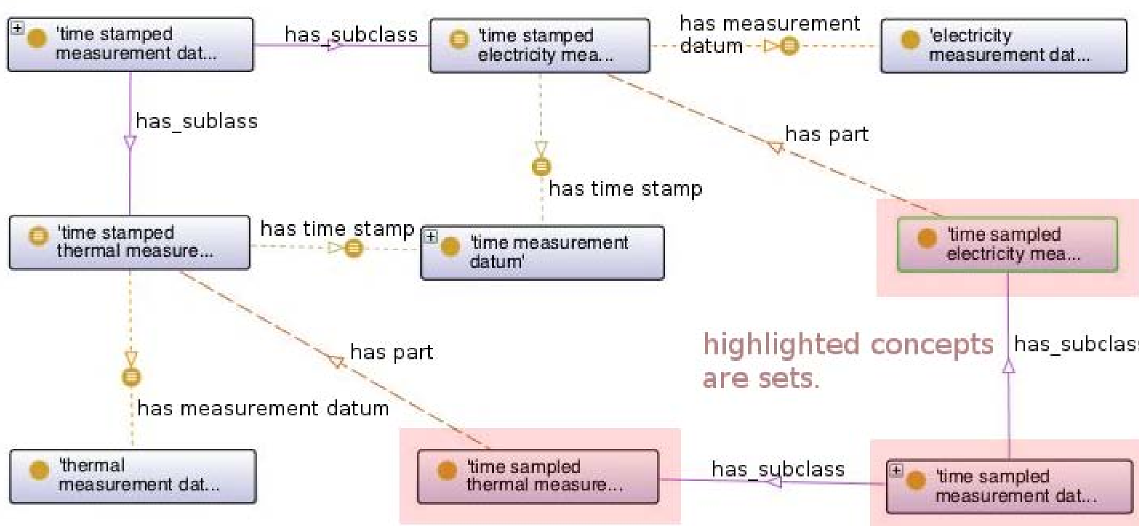

Figure 6: Semantic Relationships for Energy-Related Time Series Data

apples with oranges (as electricity measurements differ from thermal measurements), it should be noted that with the power of formal ontologies, such as OEI, algorithms are able to infer that both concepts are essentially "fruit" and take advantage of that semantic information. This is facilitated through the reasoning process for ontologies. In the example above, both *time sampled electricity measurement data set* and *time sampled thermal measurement data set* are specified in such a way that general time-series algorithms can be applied to instances of either (or both) concept for the purpose of gathering statistics. If a deeper analysis is required, the necessary relationships for each respective term (i.e., the relationships and inferences surrounding thermal measurements and electricity measurements) are also present. Take the example shown in Figure 7. Any combination of rows or columns in the conceptual division depicted can be targeted for the purposes of analysis.

High level analysis can be performed on generic time-series data or specific analysis can be done on thermal or electricity measurement data. If time-series weather data is available, it can also be

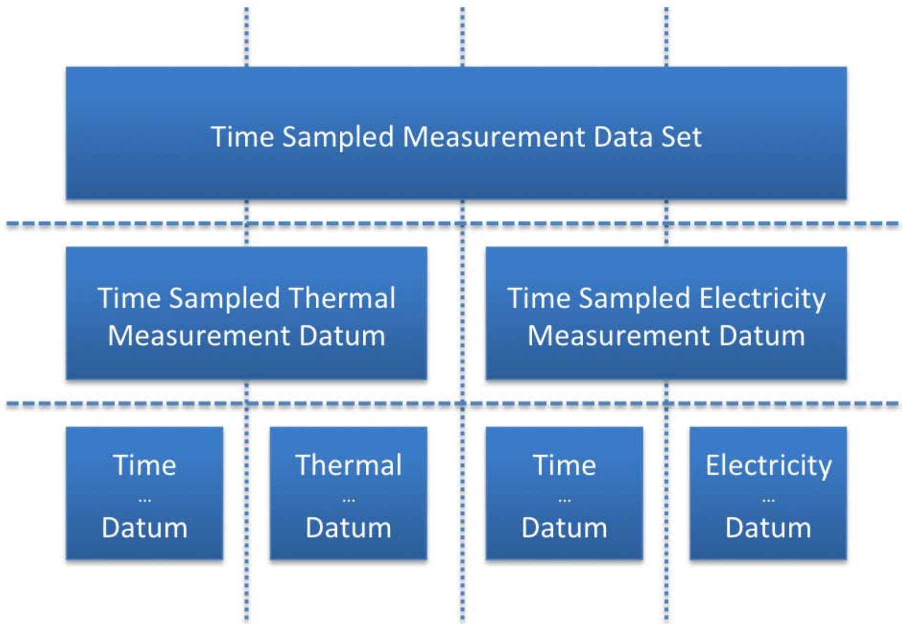

Figure 7: Axes for Multi-Level Analysis

annotated with ontological concepts. Any system that is able to infer the relationships would be able to perform a higher level statistical analysis (e.g., a correlational analysis) on the combined data sets. The ability to perform multi-dimensional analysis such as this makes the system more scalable. Code written for programs that perform such analyses can target both high level (general) and low level (specific concepts) thus increasing the ease of implementation and breadth of usefulness.

Related Work

There are also other, existing ontologies in the energy and Power Systems space from which we can use to build up our knowledge base. However, in general, these other ontologies do not take advantage of formal alignment techniques and are rarely rooted in the BFO.

The system architecture proposed in Han et al. (2011) includes an ontology model for efficient building energy management systems with concepts for sensors, equipment, zones, buildings, equipment action, zone evaluation, etc. It also includes a good description of how to build more inference rules into the ontology reasoning process via simulation.

In Keirstead and van Dam (2010), two ontologies for agent-based modeling of energy systems, the ontology for Socio-Technical Systems (STS) and the Synthetic City (SynCity) ontology for urban energy systems, are compared. The first ontology, STS, takes a network approach to cross-domain policy (i.e., not just energy-related) modeling Van Der Sanden and Van Dam (2010). It has been used to develop multiple models related to energy policies and would probably serve as a good source for some policy-related concepts. The second ontology, SynCity, is interesting because it provides three major components for modeling purposes: a mixed-integer linear programming

(MILP) optimization model for housing layouts, an agent-based energy demand model and a MILP optimization model for combining the other two models Keirstead et al. (2010). SynCity serves more as a knowledge base (i.e., a collection of instances of ontological concepts) than just a schema for concepts and does not take advantage of many of the reasoning capabilities and consistency checks available for ontologies. Neither of these two ontologies use or share a formal high level ontology, however, they do seem like another worthwhile source of concepts related to the EI domain.

In Daouadji et al. (2010), an ontology for the Information and Communication Technologies (ICT) domain that is related to energy consumption is proposed. The novelty of this ontology is the conceptual difference between "Green Energy" (e.g., solar, wind, etc.) and "Dirty Energy" (e.g., natural gas, heating, etc.). Incorporating these kinds of relationships with the concepts being added to OEI may be worth investigating. Again, this ontology does not extend any kind of formal high level ontology to root its concepts. It is also only designed in RDF and not Web Ontology Language (OWL) and therefore lacks some of the reasoning capabilities facilitated through the use of OWL ontologies.

Conclusions & Future Work

Ontologies allow for information systems to integrate at the conceptual level, enabling analysis tools and programs to infer the meaning of data without it being hard-coded. Formal ontologies present high-level semantic concepts and their relationships in a standard (and formal) way in order to facilitate reuse by different domain-specific ontologies. While many fields related to energy systems have adopted semantic technologies to develop shared vocabularies and ontologies, these have not been rooted in a formal ontology that can be shared with other domains. This paper presented the Ontology for Energy Investigations (OEI), an ontology targeted at providing a core structure that helps relate domain-specific ontologies and vocabularies dealing with energy systems back to the Basic Formal Ontology (BFO), taking advantage of both general, scientific and energy-related ontological terms and relationships already developed in the biomedical domain in the Ontology for Biomedical Investigations (OBI).

This paper described the development of OEI, demonstrating the utility rooting an energy-system related ontology in a shared formal ontology. Additionally, examples of how the National Renewable Energy Laboratory (NREL) is integrating OEI in its own data processing systems to encourage generative development and sharing of its energy models and software was provided as well facilitate multi-level analysis of the data contained in those models.

Future Work

Pending certain organizational efforts by the OBI Developers Group to finalize a stable OBI core that can be the basis for non-biomedical applications such as OEI, we will continue to develop, maintain and extend OEI in order to facilitate the benefits of a unified formal ontology for the EI domain. In addition, we will explore further application-specific uses for OEI in the EI domain.

How to Contribute to the OEI project

As the development of OEI is an ongoing process, changes happen on a daily basis, with special care taken to maintain a consistent ontology release. The latest development version of OEI is available on-line at `http://code.google.com/p/oei-ontology/`. If you are interested in getting involved with the OEI project, please visit the site just mentioned for more information on how to participate. There are also two mailing lists available: you can subscribe to the developers list at `http://groups.google.com/group/oei-developers` and the users list at `http://groups.google.com/group/oei-users`.

Acknowledgments

The authors would like to thank many of the groups at the National Renewable Energy Laboratory (NREL) as well as the OBI Developers Group for their thoughtful considerations and contributions throughout the genesis and realization of this project. Specifically, we would like thank Ben Kroposki[4] and Stuart MacMillan[5] for their suggestions, guidance, and inspiration throughout this research process. Also, this work is based upon work supported by the 2012 NREL Student Research Participation Program.

References

Ambite, J., Arens, Y., Gravano, L., Hatzivassiloglou, V., Hovy, E., Klavans, J., Philpot, A., Ramachandran, U., Sandhaus, J., Singla, A., et al. (2000). Simplifying Data Access: The Energy Data Collection (EDC) Project. In *Proceedings of the 2000 Annual National Conference on Digital Government Research*, pages 1–11. Digital Government Society of North America.

Arp, R. and Smith, B. (2008). Function, Role, and Disposition in Basic Formal Ontology. *Nature Preceedings*, pages 1–4.

Bauer, F., Recheis, D., and Kaltenböck, M. (2011). data.reegle.info–A New Key Portal for Open Energy Data. *Environmental Software Systems. Frameworks of eEnvironment*, pages 189–194.

Benjamin, P., Patki, M., and Mayer, R. (2006). Using Ontologies for Simulation Modeling. In *Proceedings of the 38th Winter Simulation Conference*, pages 1151–1159. IEEE.

Bock, J., Haase, P., Ji, Q., and Volz, R. (2008). Benchmarking OWL Reasoners. In *Proceedings of the ARea2008 Workshop*, Tenerife, Spain.

Brinkman, R., Courtot, M., Derom, D., Fostel, J., He, Y., Lord, P., Malone, J., Parkinson, H., Peters, B., Rocca-Serra, P., et al. (2010). Modeling Biomedical Experimental Processes with OBI. *Biomedical Semantics*, 1(Suppl 1):S7.

[4]Ben Kroposki leads NRELs strategic research in Energy Systems Integration.

[5]Stuart MacMillan is a Consulting Professor at Stanford University and Chief Scientist in Energy Informatics at NREL.

Brodt-Giles, D. (2012). OpenEI An Open Energy Data and Information Exchange for International Audiences. In *Proceedings of the 2012 World Renewable Energy Forum*. International Solar Energy Society.

Bussler, C., Fensel, D., and Maedche, A. (2002). A Conceptual Architecture for Semantic Web Enabled Web Services. *ACM Sigmod Record*, 31(4):24–29.

Chinnici, R., Moreau, J., Ryman, A., and Weerawarana, S. (2009). Web Services Description Language (WSDL) Version 2.0 Part 1: Core Language. Technical report, W3C. `http://www.w3.org/TR/wsdl20/` (Accessed June 25, 2012).

Daouadji, A., Nguyen, K., Lemay, M., and Cheriet, M. (2010). Ontology-based Resource Description and Discovery Framework for Low Carbon Grid Networks. In *First IEEE International Conference on Smart Grid Communications*, pages 477–482. IEEE.

Dhamanaskar, A., Cotterell, M., Zheng, J., Kissinger, J., Stoeckert Jr, C., and Miller, J. A. (2012). Suggestions for Galaxy Workflow Design Using Semantically Annotated Services. In *Proceedings of the 7th International Conference on Formal Ontology in Information Systems*.

Farrell, J. and Lausen, H. (2009). Semantic Annotations for WSDL and XML Schema. Technical report, W3C. `http://www.w3.org/TR/sawsdl/` (Accessed June 25, 2012).

Gennari, J., Musen, M., Fergerson, R., Grosso, W., Crubézy, M., Eriksson, H., Noy, N., and Tu, S. (2003). The evolution of Protégé: An Environment for Knowledge-Based Systems Development. *International Journal of Human-Computer Studies*, 58(1):89–123.

Grau, B., Horrocks, I., Motik, B., Parsia, B., Patel-Schneider, P., and Sattler, U. (2008). OWL 2: The Next Step for OWL. *Web Semantics: Science, Services and Agents on the World Wide Web*, 6(4):309–322.

Hadley, M. (2009). Web Application Description Language. Technical report, W3C. `http://www.w3.org/Submission/wadl/` (Accessed June 25, 2012).

Han, J., Jeong, Y., and Lee, I. (2011). Efficient Building Energy Management System Based on Ontology, Inference Rules, and Simulation. In *Proceeindgs of the Internaional Conference on Intelligent Building and Management*, pages 295–299.

Heath, T. and Bizer, C. (2011). Linked data: Evolving the Web into a Global Data Space. *Synthesis Lectures on the Semantic Web: Theory and Technology*, 1(1):1–136.

Hendler, J., Holm, J., Musialek, C., and Thomas, G. (2012). US Government Linked Open Data: semantic.data.gov. *IEEE Intelligent Systems*.

Horridge, M., Drummond, N., Goodwin, J., Rector, A., Stevens, R., and Wang, H. (2006). The Manchester OWL Syntax. *OWL: Experiences and Directions*, pages 10–11.

Hovy, E. (2003). Using an Ontology to Simplify Data Access. *Communications of the ACM*, 46(1):47–49.

Hu, Z., Hung, J., Wang, Y., Chang, Y., Huang, C., Huyck, M., and DeLisi, C. (2009). VisANT 3.5: Multi-Scale Network Visualization, Analysis and Inference based on the Gene Ontology. *Nucleic Acids Research*, 37(suppl 2):W115–W121.

Keirstead, J., Samsatli, N., and Shah, N. (2010). SynCity: An Integrated Tool Kit for Urban Energy Systems Modelling. *Energy Efficient Cities: Assessment Tools and Benchmarking Practices*, page 21.

Keirstead, J. and van Dam, K. (2010). A Comparison of Two Ontologies for Agent-Based Modelling of Energy Systems. In *Proceedings of the First International Workshop on Agent Technologies for Energy Systems*, pages 21–28.

Kopeckỳ, J., Vitvar, T., Bournez, C., and Farrell, J. (2007). SAWSDL: Semantic Annotations for WSDL and XML Schema. *Internet Computing, IEEE*, pages 60–67.

Krötzsch, M., Vrandečić, D., and Völkel, M. (2006). Semantic MediaWiki. In *Proceedings of the 2006 International Semantic Web Coference*, pages 935–942. Springer.

McArthur, S., Davidson, E., Catterson, V., Dimeas, A., Hatziargyriou, N., Ponci, F., and Funabashi, T. (2007). Multi-Agent Systems for Power Engineering Applications–Part II: Technologies, Standards, and Tools for Building Multi-agent Systems. *Power Systems, IEEE Transactions on*, 22(4):1753 –1759.

McIlraith, S., Son, T., and Zeng, H. (2001). Semantic Web Services. *Intelligent Systems, IEEE*, 16(2):46–53.

Miller, J. A., Baramidze, G., Sheth, A., and Fishwick, P. (2004). Investigating Ontologies for Simulation Modeling. In *Proceeding of the 37th Annual Simulation Symposium*, pages 55–63. IEEE.

Pinto, H., Gómez-Pérez, A., and Martins, J. (1999). Some Issues on Ontology Integration. In *Proceedings of the IJCAI99 Workshop on Ontologies and Problem Solving Methods: Lessons Learned and Future Trends*, pages 7.1–7.12, Stockholm, Sweden.

Pinto, H. and Martins, J. (2004). Ontologies: How can They be Built? *Knowledge and Information Systems*, 6(4):441–464.

Pisanelli, D. (2004). Biodynamic Ontology: Applying BFO in the Biomedical Domain. *Ontologies in Medicine*, 102:20.

Shearer, R., Motik, B., and Horrocks, I. (2008). HermiT: A Highly-Efficient OWL Reasoner. In *Proceedings of the 5th International Workshop on OWL: Experiences and Directions (OWLED 2008)*, pages 26–27.

Shi, W. (2010). Renewable Energy: Finding Solutions for a Greener Tomorrow. *Reviews in Environmental Science and Biotechnology*, 9:35–37.

Silver, G., Miller, J. A., Hybinette, M., Baramidze, G., and York, W. S. (2011). DeMO: An Ontology for Discrete-event Modeling and Simulation. *SIMULATION: Transactions of The Society for Modeling and Simulation International (SIM)*, 87(9):747–773.

Sirin, E., Parsia, B., Grau, B., Kalyanpur, A., and Katz, Y. (2007). Pellet: A Practical OWL-DL Reasoner. *Web Semantics: Science, Services and Agents on the World Wide Web*, 5(2):51–53.

Smith, B., Ashburner, M., Rosse, C., Bard, J., Bug, W., Ceusters, W., Goldberg, L., Eilbeck, K., Ireland, A., Mungall, C., et al. (2007). The OBO Foundry: Coordinated Evolution of Ontologies to Support Biomedical Data Integration. *Nature Biotechnology*, 25(11):1251–1255.

Spear, A. (2006). Ontology for the Twenty First Century: An Introduction with Recommendations. *Saarbrucken, Germany: Institute for Formal Ontology and Medical Information Science*, pages 1–132.

Van Der Sanden, M. and Van Dam, K. (2010). Towards an Ontology of Consumer Acceptance in Socio-Technical Energy Systems. In *Proceedings of the Third International Conference on Infrastructure Systems and Services: Next Generation Infrastructure Systems for Eco-Cities*, pages 1–6. IEEE.

Verma, K. and Sheth, A. (2007). Semantically Annotating a Web Service. *Internet Computing, IEEE*, 11(2):83–85.

W3C (2008). SPARQL Query Language for RDF. Technical report, W3C. `http://www.w3.org/TR/rdf-sparql-query/` (Accessed June 10, 2012).

W3C (2009). RDF Primer. Technical report, W3C. `http://www.w3.org/TR/rdf-primer/` (Accessed June 10, 2012).

W3C OWL Working Group (2009a). OWL 2 Web Ontology Language Direct Semantics. Technical report, W3C. `http://www.w3.org/TR/owl2-direct-semantics/` (Accessed June 10, 2012).

W3C OWL Working Group (2009b). OWL 2 Web Ontology Language Document Overview. Technical report, W3C. `http://www.w3.org/TR/owl2-overview/` (Accessed June 10, 2012).

W3C OWL Working Group (2009c). OWL 2 Web Ontology Language RDF-Based Semantics. Technical report, W3C. `http://www.w3.org/TR/owl2-rdf-based-semantics/` (Accessed June 10, 2012).

Wang, R., Guttula, C., Panahiazar, M., Yousaf, H., Miller, J. A., Kraemer, E., and Kissinger, J. (2011). Web Service Composition using Service Suggestions. In *Proceedings fo the 2011 IEEE World Congress on Services*, pages 482–489. IEEE.

Wang, X., Zhang, D., Gu, T., and Pung, H. (2004). Ontology Based Context Modeling and Reasoning using OWL. In *Proceedings of the Second IEEE Annual Conference on Pervasive Computing and Communications Workshops*, pages 18–22. IEEE.

Watson, R. T. and Boudreau, M.-C. (2011). *Energy Informatics*. Green ePress, First edition.

Young, K., Reber, T., and Witherbee, K. (2012). Hydrothermal exploration best practices and geothermal knowledge exchange on openei. In *Proceedings of the Thirty-Seventh Workshop on Geothermal Reservoir Engineering*.

Zhao, H., Zhang, S., and Zhao, J. (2012). Research of Using Protege to Build Ontology. In *Proceedings of the 11th International Conference on Computer and Information Science (ICIS)*, pages 697–700. IEEE, ACIS.

Zheng, J., Stoyanovich, J., Manduchi, E., Liu, J., and Stoeckert Jr, C. (2011). AnnotCompute: Annotation-Based Exploration and Meta-Analysis of Genomics Experiments. *Database: The Journal of Biological Databases and Curation*, 2011.

芽|Sprouts
Working Papers on Information Systems | ISSN 1535-6078

芽|Sprouts