Final Report

For the

Wind Turbine Analysis Project

May 2018

**PROJECT 15-11-03**

*Authored by:*

*J. Neil Otte, Rahul Rai, Clare Paul, and Barry Smith*

*With contributions by:*

*Ruoyu Yang, Binbin Zhang, and Munira Binti Mohd Ali*

*University at Buffalo, SUNY*

*12 Capen Hall*

*Buffalo, NY 14260-1660*

*Prepared for:*

DMDII

Contents

[**Acronyms and Abbreviations**](#_30j0zll) **3**

[**Introduction**](#_1fob9te) **4**

[**Ontologies**](#_3znysh7) **4**

[Ontologies](#_2et92p0) 4

[Tiered Ontologies](#_3dy6vkm) 4

[The WTBA Ontology](#_qsh70q) 7

[**Working with Data: Alignment and Ingestion**](#_49x2ik5) **11**

[Requirements](#_17dp8vu) 14

[**Accomplishments and Future Research**](#_sh3b04ce0p14) **14**

[Accomplishments](#_1ci93xb) 14

[Education](#_ji6e3qkbea8d) 14

[Lessons Learned](#_7gfyixty27gn) 14

[Future Research](#_y9a936rmbodn) 15

[**References**](#_3ygebqi) **17**

[**Appendix: Two Sample SPARQL Returns**](#_7x5pv96mcejt) **18**

# Acronyms and Abbreviations

|  |  |
| --- | --- |
| **Acronyms** | **Terms** |
| BFO | Basic Formal Ontology |
| CCO | Common Core Ontologies |
| CHAMP | Coordinated Holistic Alignment of Manufacturing Processes |
| PLC | The Product Life Cycle Ontologies |
| GUI | Graphical User Interface |
| IOF | Industrial Ontologies Foundry |
| IRI | International Resource Identifier |
| EMAE6 | Early Materials and Environments 2006 |
| JSON | JavaScript Object Notation |
| JSON-API | JSON Abstract Public Interface |
| R2RML | RDB to RDF Mapping Language |
| NLP | Natural Language Processing |
| OWL | Web Ontology Language |
| RDF | Resource Description Framework |
| SPARQL | SPARQL Protocol and RDF Query Language |
| URL | Uniform Resource Locator |
| URN | Uniform Resource Name |

# Introduction

Last year, the University at Buffalo, CUBRC, and Cobham Industries, carried out the Coordinated Holistic Alignment of Manufacturing Processes (CHAMP) project with the objective of enabling manufacturing organizations to overcome some of the issues caused by data heterogeneity. The CHAMP project sought to provide remedies to situations in which disparate data sources cannot be used in combination because of differences in their underlying conceptual schemas. This resulted in a suite of ontologies that are now being actively investigated for use by multiple parties within the community of ontologists working to support digital manufacturing, including the Industrial Ontologies Foundry [[1]](#footnote-1)—a NIST-led project whose goal is the creation of a shared platform for coordinated ontology creation to support interoperability of digital manufacturing software and information systems.

Following CHAMP, DMDII expressed an interest in seeing an application of ontologies within the domain of manufacturing. Out of discussions with DMDII, the members of the CHAMP team at the University at Buffalo, and Clare Paul of the Air Force Research Laboratory committed to the present project, which we have titled: Wind Turbine Analysis (WTBA). WTBA builds upon the work of CHAMP, reusing its ontologies in the building of a single application ontology necessary for the performance of a particular material analysis using a particular data set. This project will be of interest to industry professionals. This includes industry professionals who wish to immediately use the framework we have to select materials for wind turbines using our dataset and procedures, as well as the broader community, who are already using or planning to use semantic technology, and are interested in seeing how it may facilitate the discovery of data that meet a set of pre-defined constraints or parameters.

# Ontologies

## Ontologies

We follow the W3C recommendation that ontologies be published for exchange as OWL files (more precisely: as files using the OWL 2 Web Ontology Language). OWL is the current standard and there is an ecosystem of tools developed for creating ontologies within its language.

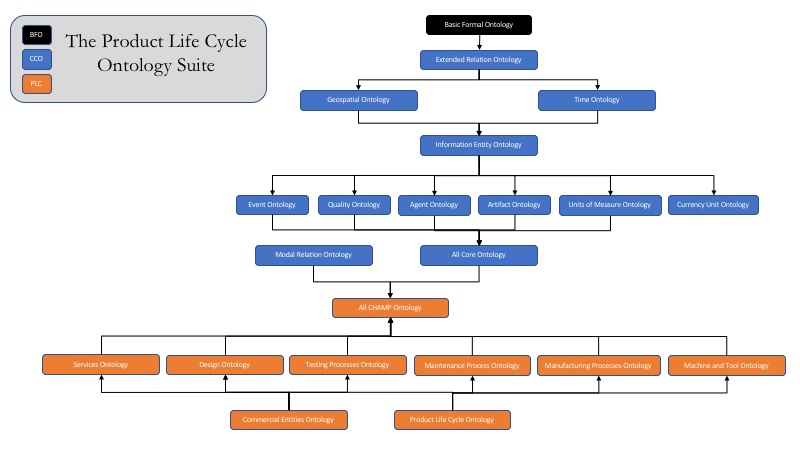
For further discussion of OWL, RDF, JSON, and other programming languages for interest to this domain, see the previously submitted CHAMP report.[[2]](#footnote-2)

### Tiered Ontologies

Contemporary ontology engineering pursues sustainability by working within commonly accepted ontologies that impose downward conformity upon the lower-tiered ontologies that use them. This practice allows ontologies developed independently to achieve a degree of compatibility that —facilitate integration of data which the ontologies are used to describe. This involves a distinction, discussed in detail in the CHAMP report, between Top-Level, Mid-Level, and Domain-Level ontologies.

Briefly, the level of an ontology is determined by the level of generality of the types in reality that it represents. The top-level concerns entities common across all domains, such as objects, times, spaces, processes, and qualities. We use for this purpose Basic Formal Ontology (BFO), which is the most widely used top-level ontology currently under review by the International Standards Organization to form ISO 21838-2. Mid-level ontologies represent those entities that are common across many domains of inquiry. These include currencies, units of measure, treatments of agents and the actions they perform, countries and states, information, and other very large domains. Here, our choice of mid-level ontologies is the suite of Common Core Ontologies (CCO).

Domain level ontologies are more particular still, and often invoke the interests of a particular community. Here, the Product Life-Cycle (PLC) Ontologies produced by the CHAMP project address data representation for industry, providing a presentation of the domains of the PLC including products, material, and data that participate in processes in its successive phases.



*Figure 2 The import structure among the ontologies of the Product Life Cycle suite, the Common Core Ontology suite, and Basic Formal Ontology*

The PLC ontologies are reference ontologies. That is, they are intended for aggressive re-use by multiple other ontologies developed for example by specific enterprises or to support specific software applications. However, user of these ontologies will often be required to extend them by creating application ontologies designed to address specific local needs. This may involve the creation of a small set of subclasses of existing parent classes within the PLC ontologies, or it may involve creating a seperate ontology for a software application that is not represented within the PLC, but which shares some representation with the PLC ontologies.

Application ontologies are the lowest tier of ontology; this is because they frequently contain classes that are particular to a data set, or are constructed in a manner required by some need of a particular end user. For this reason, unlike the PLC ontologies, application ontologies are often intended for one-off support of a particular function or form of analysis and are not intended for general re-use . Where application ontologies are successful, however, then they may serve as use cases for others, but they can do this effectively only if they are developed within a framework of mid-level and top-level ontologies that is widely accepted. Otherwise, when another user takes inspiration from an application ontology developed by some other party, the ontology that user will go on to create will more often than not lack interoperability with the application ontology that inspired it.

The outcomes of the present project should be assessed against this background. It demonstrates how ontologies can be used by a materials analyst to discover materials that will optimize wind turbine blade performance given certain parameters. As such, it represents a steel thread, or execution path through a computer system that can meet the business objectives of a company. In what follows, we detail how the present ontology was used to create data that can be used to perform a material selection procedure, and we also detail how such a tool could be commercially viable. These are immediate benefits of ontologies that demonstrate a clear value and application to DMDII members.

Because this is an application ontology, it is not created to be re-used by extended or adopted for integration by others; rather, it was designed for one specific task and with one specific data source in mind to enable performance of this task. It is also one of the use-cases that the Industrial Ontologies Foundry is presently reviewing as they pursue their long-term vision for ontology engineering coordination and practice.

### The WTBA Ontology

The Wind Turbine Blade Application (WTBA) ontology is a relatively small application ontology consisting of 82 classes. This set of classes is a conservative extension of the 1393 total classes available in the total suite of imported ontologies we are using. These classes are derived from the PLC ontologies, the ontologies of the CCO, as well as BFO and the RO.

The WTBA Ontology was developed to address the following use case:

*As a wind turbine blade engineer, I want to use artificial intelligence (AI) to recommend optimum composite material designs so that I can deliver maximum performance within environmental constraints.*

It includes representations of key parameters, such as:

• Key Performance Parameters – e.g. power efficiency, max deflection, tip velocity

• Environment – e.g. wind loads, gravity induced loads, moisture

• Design Requirements – e.g. sufficient durability, sufficient strength, minimize weight

• Design Approach – e.g. stiffness, strength, durability

• Test Methods – e.g. fatigue testing, monotonic testing, fiber-volume testing

• Engineered Material Product – e.g. carbon fiber, glass fiber, glass matt

• Product Function – e.g. resist mechanical loads, maintain reinforcement orientation

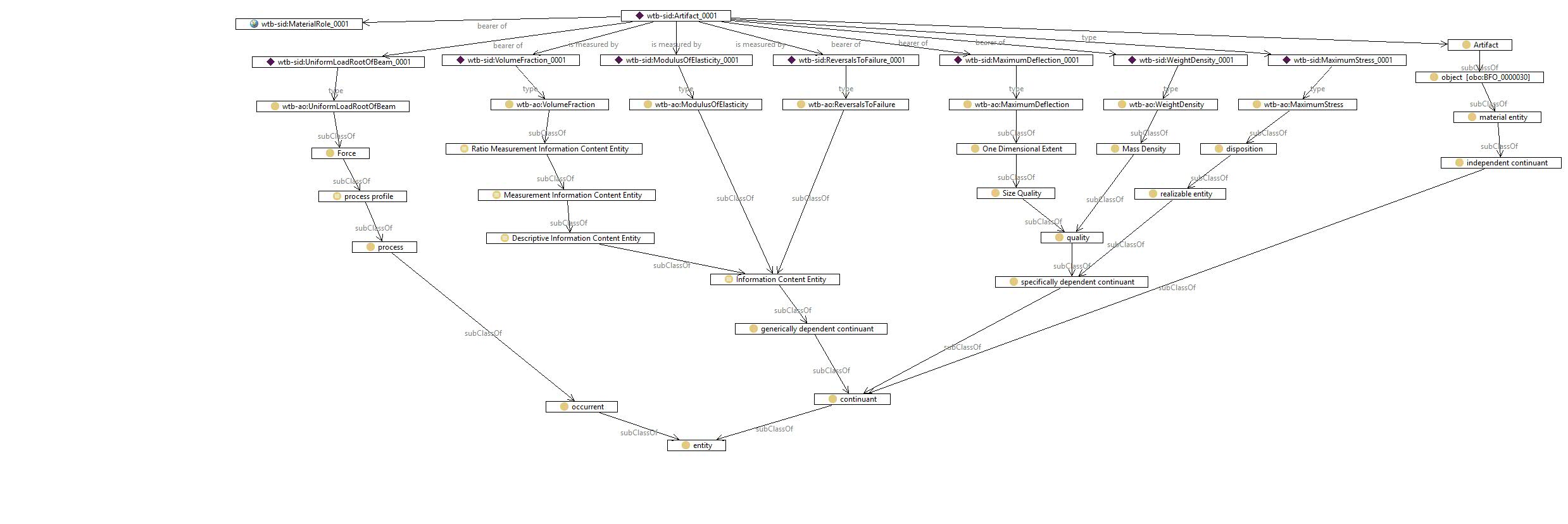
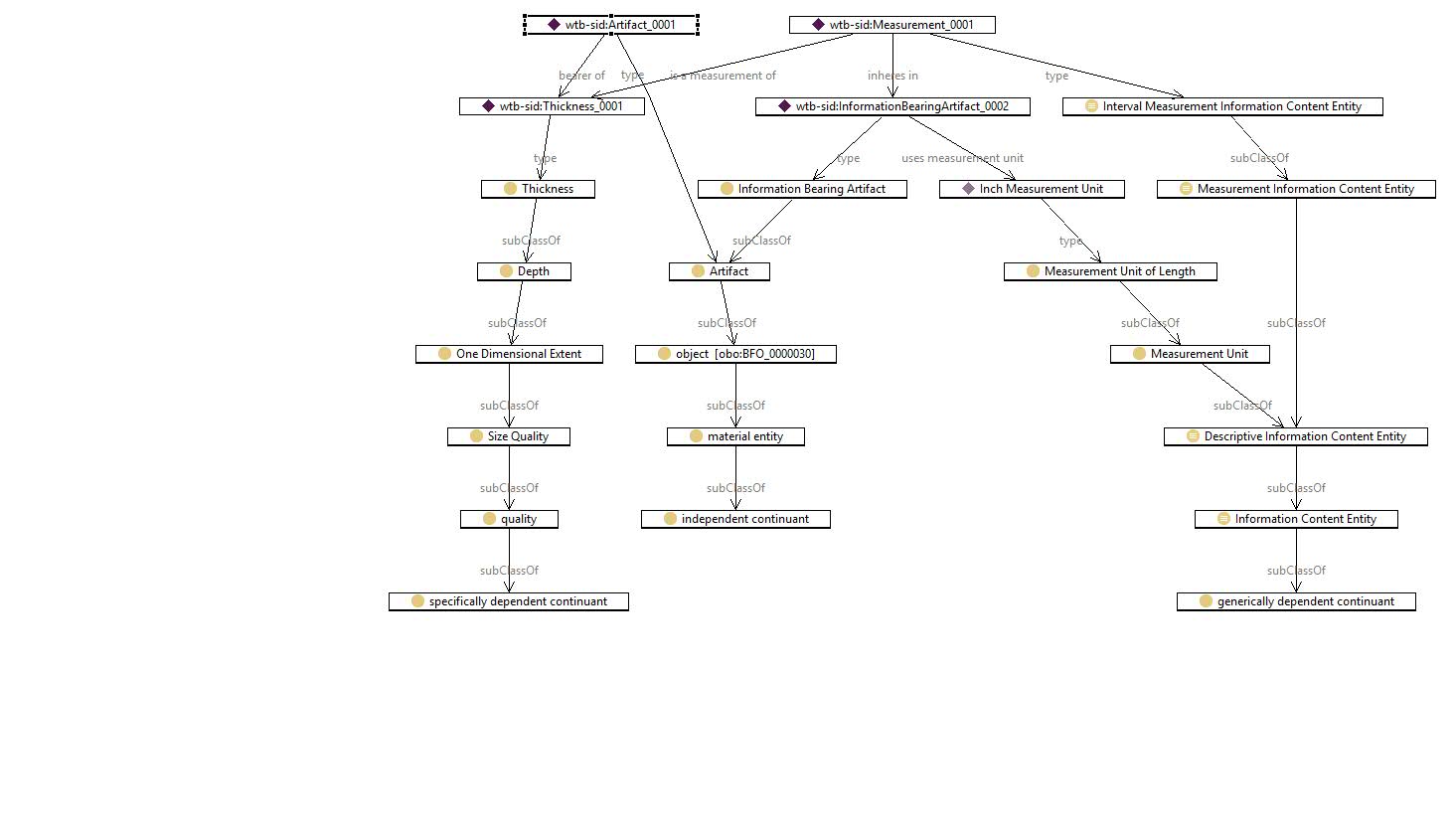
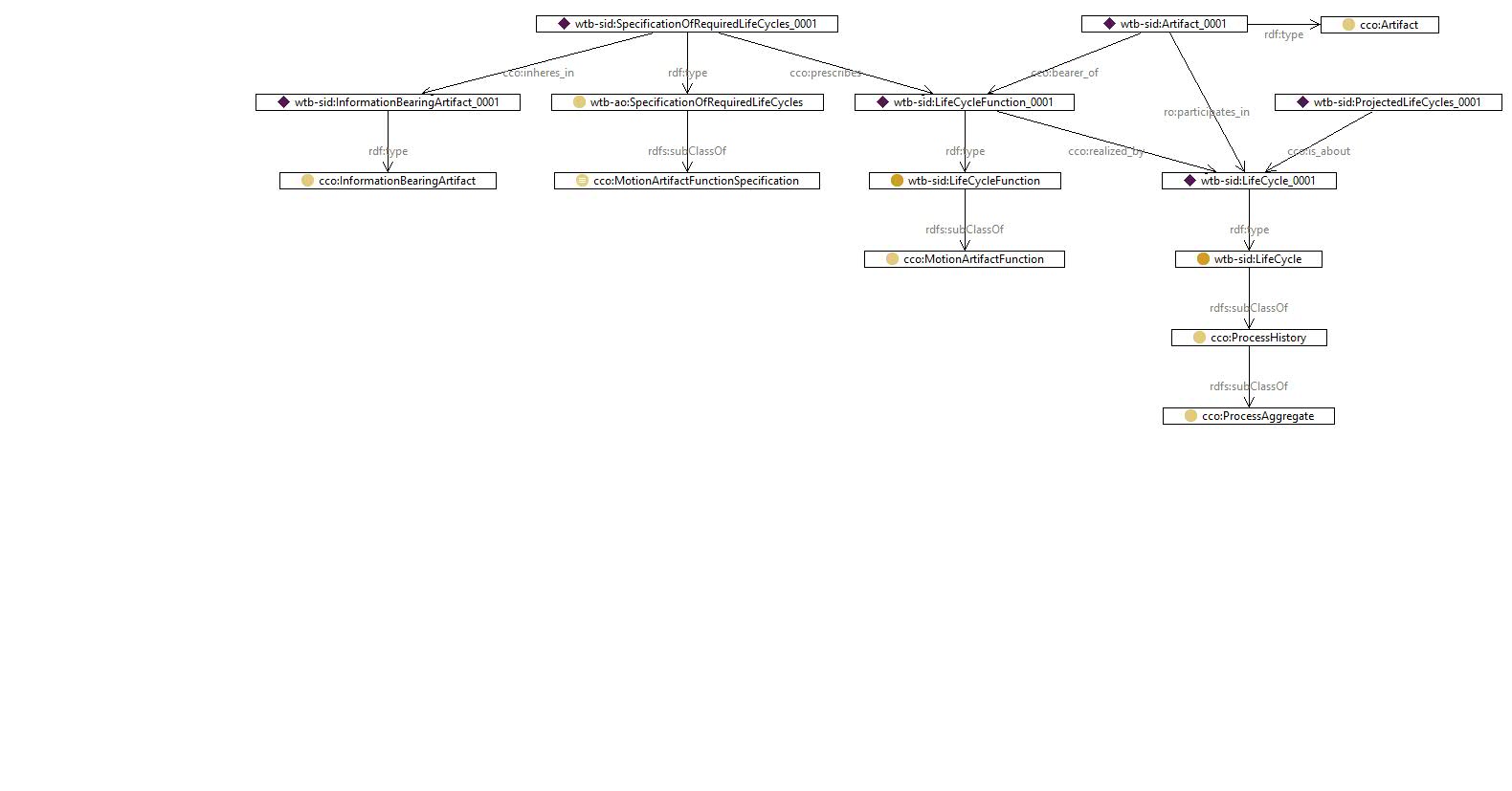
• Product Form – e.g. tow, mat, laminate

• Manufacturing Process – e.g. resin infusion, autoclave cure

*Form*

Below, we provide sample images that demonstrate the form of the knowledge graph. These images were created using TopBraid Composer—an ontology editor and data management software tool.

In each of the images, the diamonds mark instances of classes, while the circles indicate classes themselves.



# Working with Data: Alignment and Ingestion

*Dataset*

Ontologies consist not of data but rather of classes, axioms, and relations, which are used to tag data to make it discoverable, comparable, and available for certain sorts of retrieval and reasoning. The process of tagging data (or of aligning data to an ontology) can be viewed as a kind of *mapping*. A standard procedure for mapping creates two files, as follows:

1. The mapping itself, which is a software file that specifies that an element in a target data set is an instance of a class in the ontology, and that such an instance bears corresponding relations from the ontology to other instances in the data set. Mapping files may be SPARQL queries, or they may use a language, such as R2RML, which is frequently employed to create mappings between relational and graph databases.
2. The knowledge graph that is produced by the mapping process and which is commonly saved in RDF or JSON. This amounts to a set of triples that conform to the semantics of the ontology. This knowledge graph is then stored in a database where it is queried using SPARQL.

For this project, we used the dataset from an open source database for wind turbine blade materials data at Sandia National Lab at Montana State University.[[3]](#footnote-3) This materials data set includes, in particular, data pertaining to different types of reinforcement and matrix materials. In order to avoid duplicating unnecessary effort, we selected to map only those components necessary for our queries of interest, though the procedures we describe here would enable further uses of the ontology to enable other queries of interest.

The mapping itself was performed by first creating a set of dummy instances of classes in the ontology, where these instances had the form necessary for the mapping. Following this, TopBraid Composer was used to construct a series of SPARQL construct queries, which took data from the Sandia National Lab data set (in excel format) and created an RDF graphical representation of this data.

*SPARQL Queries*

Queries were then developed that conformed to the key performance parameters used by a wind turbine engineer in designing a turbine blade. These queries allowed data that took into account a variety of conditions, including material type, fiber direction, gauge size, the resin used on the material, the maximum stress applied to the material during testing, and many other variables. Such queries draw upon the computational semantics of the ontologies and SPARQL to look for data instances in the graph that satisfy the constraints of the query. This allowed, for instance, the stipulation of particular conditions, and the exploration of data over a number of scenarios involving the use within different environment of different composite materials, ply orientations, reinforcement and matrix materials, and so forth.

For example, to find the variation of the frequency (Hz) and the maximum stress (Mpa) applied to a specimen; according to the material of which it is a type; its fiber direction; and gauge section width and thickness, the following SPARQL query was used:

**Query 1.**

|  |
| --- |
| PREFIX emae06: <http://example.com/emae06#>  PREFIXxsd: <http://www.w3.org/2001/XMLSchema#>  SELECT DISTINCT ?Material ?Resin  ?ReinforcementFiber?ZeroDegreeDirectionFiber\_pct?FortyFiveDegreeDirectionFiber\_pct?NinetyDegreeDirectionFiber\_pct?OtherDegreeDirectionFiber\_pct?GaugeSectionWidth\_mm?GaugeSectionThickness\_mm?MaximumStress\_MPa?FatigueTestFrequency\_Hz  WHERE {  ?a emae06:sshUsesMaterial ?Material .  ?a emae06:sshResin ?Resin .  ?a emae06:sshZeroDegreeReinforcementFabric ?ReinforcementFiber.  ?a emae06:sshPctInZeroDegDirection ?ZeroDegreeDirectionFiber\_pct.  ?a emae06:sshPctInFortyFiveDegreeDirection ?FortyFiveDegreeDirectionFiber\_pct.  ?a emae06:sshPctInNinetyDegreeDirection ?NinetyDegreeDirectionFiber\_pct.  ?a emae06:sshPctInOtherDirection ?OtherDegreeDirectionFiber\_pct.  ?a emae06:sshGaugeSectionThickness ?GaugeSectionThickness\_mm.  ?a emae06:sshGageSectionWidth ?GaugeSectionWidth\_mm.  ?a emae06:sshMaxStress ?MaximumStress\_MPa.  ?a emae06:sshFatigueTestFrequency ?FatigueTestFrequency\_Hz .  }  ORDER BY ?Material ?Resin ?ReinforcementFiber?ZeroDegreeDirectionFiber\_pct?FortyFiveDegreeDirectionFiber\_pct?NinetyDegreeDirectionFiber\_pct?OtherDegreeDirectionFiber\_pct?GaugeSectionWidth\_mm?GaugeSectionThickness\_mm |

Note, the prefix ‘emae06’, short for the name of the MSU data set ‘Early Materials and Environments 2006’, denotes the base URI of the application ontology.

A complete list of SPARQL queries is provided at the github repository[[4]](#footnote-4) where they are available for re-use, along with the original RDF file. We list below in natural language a sample of other queries that we created:

|  |
| --- |
| **Query 1**  According to the material type, its fiber direction, and the specimen’s dimensions, find the variation in the frequency and the maximum stress applied to the specimen.  **Query 1a**: Filter the search by the material’s layer structure.  **Query 1b**: Filter the search by the material’s name[[5]](#footnote-5).  **Query 2**  According to the material type, its fiber direction, and the specimen’s dimensions, find the variation of the maximum strain percentage and the load cycle time applied to the specimen.  **Query 2a**: Filter the search by the material’s layer structure.  **Query 2b**: Filter the search by the material’s name.  **Query 3**  According to the material type, its fiber direction, and the specimen’s dimensions, find the variation of the stress ratio (min stress/max stress) and the load cycle time applied to the specimen.  **Query 3a**: Filter the search by the material’s layer structure.  **Query 3b**: Filter the search by the material’s name. |

## 

## Requirements

The mid-level ontologies in the PLC suite that were used in this project were vetted in three different ways. First, the ontologies were validated by comparing the terms in the ontologies to the terms in the dataset and confirming that there is adequate coverage in the ontology for the dataset. Second, ontology reasoners such as FaCT++ and Hermit were used to confirm the consistency[[6]](#footnote-6) of the ontology classes and axioms and conformance to semantic technology standards. Third, the ontology content was assessed for its conformity to the principles of the BFO framework.

# Accomplishments and Future Research

## Accomplishments

### Education

For one semester, the WTBA project funded three graduate students, two of whom are from the engineering school at the University of Buffalo.

In addition to the opportunities offered to students, work on the WTBA led to collaborations with a number of academic partners, who are now, with the University at Buffalo, involved actively in the Industrial Ontologies Foundry.

These institutions and their lead members include:

* INP-ENIT, University of Toulouse (Hedi Karray)
* Clemson University (Venkat Krovi)
* École polytechnique fédérale de Lausanne (Dimitris Kiritsis)
* Loughborough University, UK (Bob Young)
* National Center for Ontological Research (Kemper Lewis, Rahul Rai, and Barry Smith)
* Penn State (Timothy Simpson)
* Texas State (Farhad Ameri)
* UMass Amherst (Ian Grosse)
* University of Toronto (Michael Grüninger)

### 

### Lessons Learned

For the students attached to the project, their primary focus has been ontology engineering research and the application of design principles in building OWL ontologies using BFO principles. This project presented them with a nice transition into the realm of engineering, where the ontologies they have built are now applied to data within a particular project. This involved becoming significantly more familiar with running servers and writing SPARQL queries.

The application ontology design itself neatly falls out of the work on the PLC ontologies, which the team had spent a good deal of time perfecting. This project, indeed, underscores the value of working with mid-level ontologies like the CCO and the PLC, because working within such a framework makes very clear to the application ontology engineer how to create new classes for particular applications in a manner that may be understood well by others. We were thus able to avoid many of the basic errors that plague ontologies and data models within the realm of industry, such as those detailed by Smith (2006) concerning the ISO standard 15926 (“Lifecycle Integration of Process Plant Data Including Oil and Gas Production Facilities”).

For Clare Paul, who provided the technical expertise necessary to the completion of this project, the project provided an opportunity to become more familiar with the basic representation of measurement data provided by the Common Core Ontologies. This familiarity will be valuable in the future work to occur on the materials property ontology, whose alignment with the PLC and CCO ontologies is an ongoing effort.

## Future Research

DMDII members are clearly concerned with the commercial viability of this project. The primary steps for using the deliverables of this project are relatively simple for most technical experts.

They include:

1. Downloading and installing an RDF triple store, such as Apache Jena
2. Loading RDF files into the server
3. Running SPARQL queries in the server endpoint, or writing other queries of interest.

However, in order to be truly commercially viable, even these steps might prove off-putting to many DMDII members who are unaccustomed to running a local server and understanding SPARQL syntax.

An alternative is to develop a graphical user-interface that presents drop-down options to a user, allowing them to select what they wish to find and the constraints they wish to place on their search. Such forms can then write SPARQL queries and run these queries against a backend server without the user needing to understand SPARQL, RDF, or any of the semantic technologies involved.

One example of such a tool is OWL2TL[[7]](#footnote-7). This open, online tool presents a series of fields to users, whose inputs are used to construct a SPARQL query that queries the annotation properties of ontology files. The SPARQL query is run on the backend of the website. This would be one way of making the sort of system we have built considerably more simple and user-friendly.

# References

[1] Arp, R., Smith, B. and Spear, A., 2015. Building Ontologies with Basic Formal Ontologies.

[2] Frechette, S.P., 2011. Model based enterprise for manufacturing. Duffie, ed., Omnipress, Madison, WI

[3] Otte, J. N., Ruttenberg. A. “BFO: Basic Formal Ontology”. *Applied Ontology* [Forthcoming]

[4] Pettey, C, 2016, 5 Ugly Truths About Postmodern ERP (Gartner)<https://www.gartner.com/smarterwithgartner/5-ugly-truths-about-postmodern-erp/>

[5] Shen, W, et al, 2008, Computer supported collaborative design: Retrospective and perspective, Computers in Industry, 59/9: 855-862.

[6] Smith, B. 2006. “Against Idiosyncrasy in Ontology Development”. B. Bennett and C. Fellbaum (Eds.), *Formal Ontology and Information Systems*, (FOIS 2006), Amsterdam: IOS Press, 2006, 15-26. Preprint version available here: <http://ontology.buffalo.edu/bfo/west.pdf>

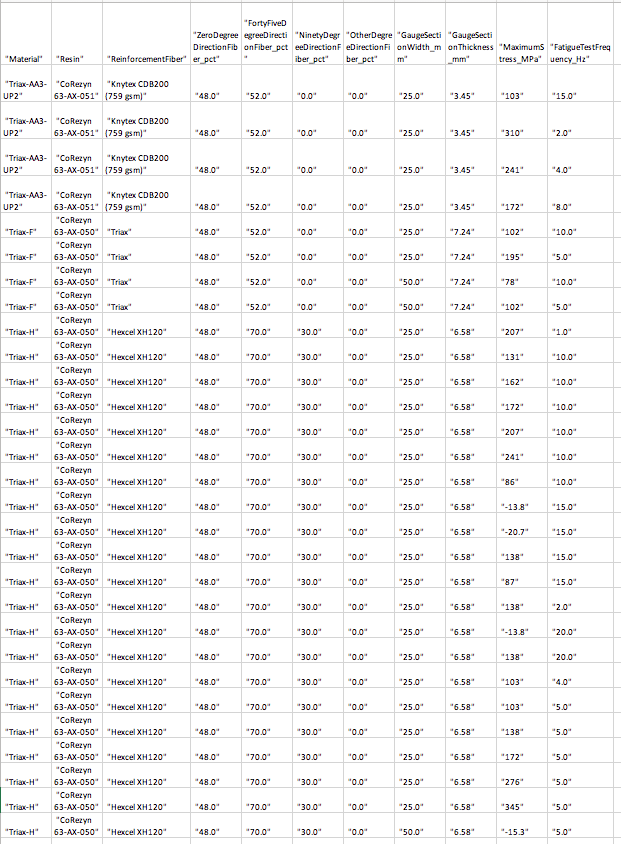
[7] Smith, B. and Ceusters, W., 2010. Ontological realism: A methodology for coordinated evolution of scientific ontologies. *Applied ontology*, *5*(3-4), pp.139-188.

[8] Smith, Barry et al. “The OBO Foundry: Coordinated Evolution of Ontologies to Support Biomedical Data Integration.” *Nature biotechnology* 25.11 (2007): 1251. *PMC*. Web. 10 Jan. 2018.

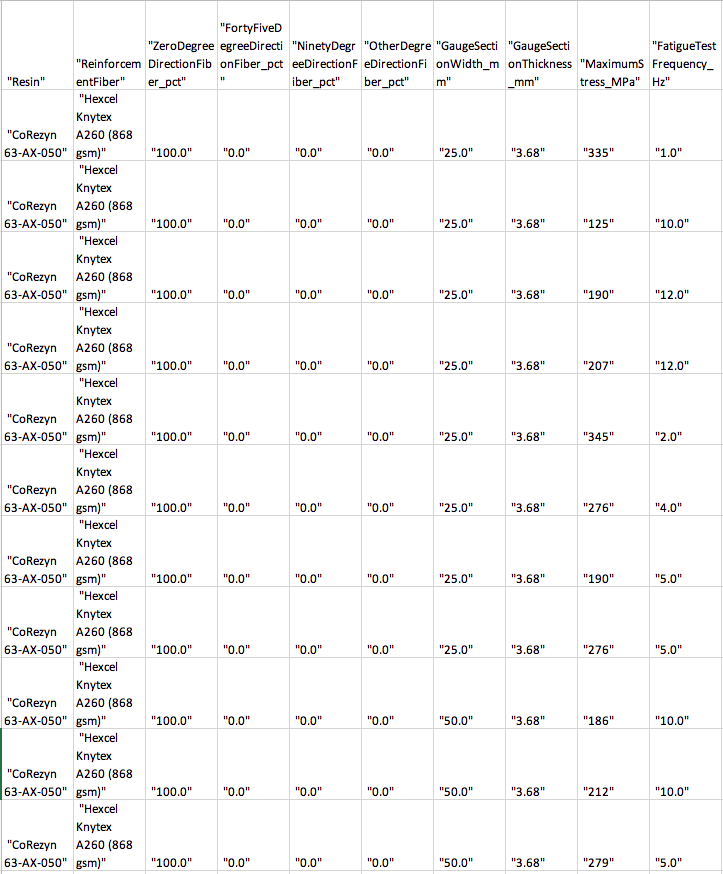
[9] Furini F, Rai R, Smith B, Colombo G, Krovi V. Development of a Manufacturing Ontology for Functionally Graded Materials. ASME. International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Volume 1B: 36th Computers and Information in Engineering Conference ():V01BT02A030. doi:10.1115/DETC2016-59964.

# Appendix: Two Sample SPARQL Returns

Query 1A Return



Query 1b Return



1. The website of the IOF may be accessed here: <https://sites.google.com/view/industrialontologies/home> [↑](#footnote-ref-1)
2. The final CHAMP report may be accessed here: https://github.com/NCOR-US/CHAMP/blob/master/CHAMP%20Base%20Year%20Final%20Technical%20Report.docx [↑](#footnote-ref-2)
3. The dataset is described and may be accessed here: http://energy.sandia.gov/energy/renewable-energy/wind-power/blade-reliability/blade-materials-and-substructures-testing/ [↑](#footnote-ref-3)
4. All files related to the Wind Turbine Project are available here: <https://github.com/NCOR-US/WindTurbineProject> [↑](#footnote-ref-4)
5. Note: We include in this report the returned results for queries 1a and 1b (see Appendix). For a full list of all returns, see the project repository on github in fn 4. [↑](#footnote-ref-5)
6. For an explanation of consistency in the context of ontologies, see the final CHAMP report. [↑](#footnote-ref-6)
7. OWL2TL may be accessed here: https://owl2tl.com. [↑](#footnote-ref-7)