An Ontology for a Polymer Nanocomposite Community Data Resource

Vipula Rawte
Department of Computer Science
Rensselaer Polytechnic Institute
Troy, New York 12180
rawtev@rpi.edu

James McCusker
Department of Computer Science
Rensselaer Polytechnic Institute
Troy, New York 12180
mccusj2@rpi.edu

He Zhao
Department of Mechanical Engineering
Northwestern University
Evanston, Illinois 60201
hezhao2012@u.northwestern.edu

L. Catherine Brinson
Department of Materials Science and Engineering
Northwestern University
Evanston, Illinois 60201
cbrinson@northwestern.edu

Wei Chen
Department of Mechanical Engineering
Northwestern University
Evanston, Illinois 60201
weichen@northwestern.edu

Linda Schadler
Department of Materials Science and Engineering
Rensselaer Polytechnic Institute
Troy, New York 12180
schadl@rpi.edu

Deborah L. McGuinness
Departments of Computer, Cognitive, and Web Sciences
Rensselaer Polytechnic Institute
Troy, New York 12180
dlm@cs.rpi.edu

ABSTRACT

Interest in polymer nanocomposites has been increasing because of their unique properties. Since the field draws on a wide range of disciplines, a data resource aimed at adequately supporting the field needs to include vocabulary from many disciplines. We are building an ontology that aims to connect to relevant vocabularies and include well-specified interconnections between those vocabularies. Coverage is determined by a set of use cases. We will describe our ontology along with the growing data resource that it supports and introduce how our ontology-based approach can enable collaborative community research in polymer nanocomposites. Further, we mention how this approach can be leveraged more broadly in material science.

CCS CONCEPTS

•**Information systems** → **Web Ontology Language (OWL);** *Extensible Markup Language (XML);* **Ontologies; Semantic web description languages;**

KEYWORDS

knowledge graph; ontology; semantic data resource; microstructure; polymer nanocomposite; visualization; nanofiller dispersion; interfacial area

WebSci '17, June 25-28 2017 Troy, NY, USA,
 978-1-4503-4896-6/17/06.
DOI: 10.1145/3091478.3098866

1 INTRODUCTION

We are building an open access, user-friendly, living, growing, data resource for the polymer nanocomposites community called NanoMine. We aim for it to be scalable and enable improved understanding of key processing - structure - property (p-s-p) relationships. We aim for it to be used by domain scientists to facilitate faster nanocomposite design and insertion into advanced applications [7]. By bringing together data that is scattered throughout public literature and private files and by creating a protocol for recording and tagging data, this resource has the potential to be an unprecedented compilation of accessible information.. We will introduce our ontology along with the provenance annotation strategy and briefly describe the ontology and provenance usage in building and accessing a knowledge graph aimed at helping domain scientists explore information relevant to designing new polymer nanocomposite materials.

Members of the materials science community along with the National Institute of Standards and Technology have produced NIST's Materials Data Curation System (MDCS) that captures, shares, and transforms materials data into an XML-based structured format. This supports data organization using user-selected XML Schema encoded templates that are used to build data entry forms. The MDCS approach is web based using the full web stack of tools including Python, Django, MongoDB, REST API, XML, and SPARQL queries but lacks visualization of the collected data for better analysis and understanding [1].

The initial version of NanoMine used only an XML schema to address the problem of identifying relationships between processing, structure, and properties [8]. Our version of NanoMine leverages

deeper semantic relationships encoded in an OWL ontology and supports enhanced semantic queries.

2 RESULTS AND DISCUSSION

Our NanoMine ontology is created using the XML Schema and also domain knowledge about polymer nanocomposites. We are using other vocabularies like Semanticscience Integrated Ontology (SIO) [2], Nanopublications Vocabulary [5], and also planning for PubChem [3] and Units Ontology [4]. The data is collected, represented and stored in the form of nanopublications.

The ontology currently includes the basic concepts of characterization methods, materials, microstructure, processing methods, properties and publications and is being expanded based on use cases.

The ontology supports structured queries that go well beyond the traditional keyword search approach. Our community wanted to ask questions such as "What is the highest reported T_g change in a polymer nanocomposite? Or in a PMMA-based nanocomposite?" Prior to the XML-based Nanomine effort, the community could leverage web searches such as "polymer nanocomposite" and possibly refine those using "glass transition" or by "thermal property" but the answers to these queries may miss many relevant papers and the process was very slow and painful [6].

The previous XML-based NanoMine was able to encode some relatively straightforward simple relationships for nanomaterials but had very limited search capabilities and no visualization options.

We can now support queries using any parameter in the ontology and display answers textually or graphically and support improved exploration concerning different factors related to elements in the query (e.g. T_g) and their sources. We support this search using the Nanomaterials Knowledge Graph, built using our RPI Tetherless World Satoru Knowledge Graph Project. The ontology is used to connect related terms together, to provide support in hierarchically-based faceted browsing and ontology-enabled visualizations.

Currently, NanoMine is hosted on Northwestern University's server at http://nanomine.northwestern.edu:8001/.

Provenance is encoded using nanopublications. Figure 1 displays an assertion and Figure 2 shows the related publication from which it was obtained.

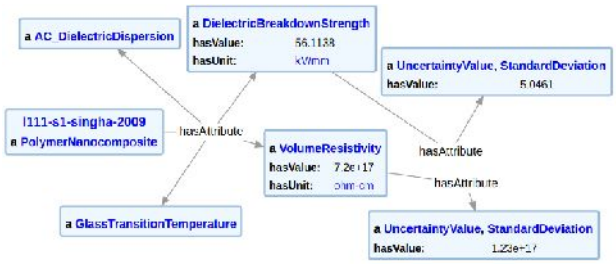

Figure 1: Assertion.

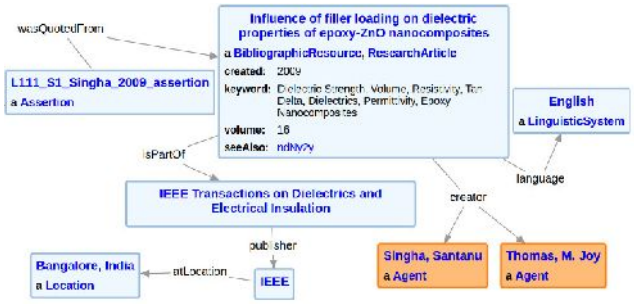

Figure 2: Publication Information.

The foundational infrastructure supports provenance encoding and the domain curation is handled by our materials collaborators at Northwestern and RPI.

3 CONCLUSIONS AND FUTURE WORK

Our provenance-aware semantic data resource is currently supporting a range of material scientists in encoding content related initially to processing-structure-property relationships. Our goal is to enhance computer and human understanding of polymer nanocomposites, thus enabling next generation collaborative research. The ontology and the knowledge graph that uses it are used to support visualizations including a customizable dashboard aimed at nanomaterials researchers. The work is in early stages, however, is currently available for collaboration and reuse.

ACKNOWLEDGMENTS

This work is a collaboration between Rensselaer Polytechnic Institute and Northwestern University and is funded through NSF ACI-1640840.

REFERENCES

[1] Alden Dima, Sunil Bhaskarla, Chandler Becker, Mary Brady, Carelyn Campbell, Philippe Dessauw, Robert Hanisch, Ursula Kattner, Kenneth Kroenlein, Marcus Newrock, et al. 2016. Informatics Infrastructure for the Materials Genome Initiative. *JOM* 68, 8 (2016), 2053–2064.

[2] Michel Dumontier, Christopher JO Baker, Joachim Baran, Alison Callahan, Leonid Chepelev, José Cruz-Toledo, Nicholas R Del Rio, Geraint Duck, Laura I Furlong, Nichealla Keath, et al. 2014. The Semanticscience Integrated Ontology (SIO) for biomedical research and knowledge discovery. *Journal of biomedical semantics* 5, 1 (2014), 14.

[3] Gang Fu, Colin Batchelor, Michel Dumontier, Janna Hastings, Egon Willighagen, and Evan Bolton. 2015. PubChemRDF: towards the semantic annotation of PubChem compound and substance databases. *Journal of cheminformatics* 7, 1 (2015), 34.

[4] Georgios V Gkoutos, Paul N Schofield, and Robert Hoehndorf. 2012. The Units Ontology: a tool for integrating units of measurement in science. *Database* 2012 (2012), bas033.

[5] Paul Groth, Andrew Gibson, and Jan Velterop. 2010. The anatomy of a nanopublication. *Information Services & Use* 30, 1-2 (2010), 51–56.

[6] Bharath Natarajan, Yang Li, Hua Deng, L Catherine Brinson, and Linda S Schadler. 2013. Effect of interfacial energetics on dispersion and glass transition temperature in polymer nanocomposites. *Macromolecules* 46, 7 (2013), 2833–2841.

[7] Richard A Vaia and John F Maguire. 2007. Polymer nanocomposites with prescribed morphology: going beyond nanoparticle-filled polymers. *Chemistry of Materials* 19, 11 (2007), 2736–2751.

[8] He Zhao, Xiaolin Li, Yichi Zhang, Linda S Schadler, Wei Chen, and L Catherine Brinson. 2016. Perspective: NanoMine: A material genome approach for polymer nanocomposites analysis and design. *APL Materials* 4, 5 (2016), 053204.