Creating and Visualizing the Materials Science Knowledge Graph with Whyis *

Jamie McCusker[1 [0000-0003-1085-6059]], Michael Deagen[2 [0000-0002-8034-0667]], Tolulomo Fateye[3 [0000-0002-2061-2261]], Anya Wallace[4 [0000-0003-3123-2122]], Sabbir M. Rashid[1 [0000-0002-4162-8334]], and Deborah L. McGuinness[1 [0000-0001-7037-4567]]

[1] Department of Computer Science, Rensselaer Polytechnic Institute, Troy, NY, US
{mccusj2,rashis}@rpi.edu, dlm@cs.rpi.edu

[2] College of Engineering and Mathematical Sciences, University of Vermont, Burlington, VT, US Michael.Deagen@uvm.edu

[3] Department of Mechanical Engineering and Materials Science, Duke University, Durham, NC, US tolulomo.fateye@duke.edu

[4] Department of Mechanical Engineering and Applied Physics, California Institute of Technology, Pasadena, CA, US awallace@caltech.edu

Abstract. The NanoMine knowledge graph is being expanded to support a broader range of materials, now including metamaterials, and is now called MaterialsMine. In the process, we have added new knowledge curation and visualization capabilities to Whyis, the framework that MaterialsMine (and NanoMine before it) is built on. This demonstration will show how we use Whyis to support user-provided data uploads that conform to the Dataset Catalog standard, use DOI and ORCiD to populate dataset metadata, curate data files into knowledge graph fragments using Semantic Data Dictionaries, and allow domain scientists to visualize and explore the graph using SPARQL, Vega-Lite, and Data Voyager.

Keywords: Semantic Science · Knowledge Graphs · Knowledge Engineering.

1 Introduction

We will demonstrate the MaterialsMine knowledge graph and how it was developed using the Whyis Knowledge Graph Framework [4]. Materials science studies the interrelationships between processing and properties of materials by analyzing material structure at multiple length scales. The MaterialsMine knowledge graph builds on the NanoMine knowledge graph [3] by expanding into datasets for mechanical and acoustic metamaterials and generalizing the ability to curate data from materials science. Metamaterials are materials that have their properties modified through their structures, resulting in radically different accoustic,

 The authors gratefully acknowledge support of the NSF CSSI program (OAC-1835677).

optical, and mechanical properties. The "invisibility cloak"[7] is probably the most well-known example for optical metamaterials, but even ancient techniques, like knitting [2], allowed the development of stretch fabrics centuries before the development of stretch fibers. In MetaMine, the metamaterials component of MaterialsMine, we have developed a data and knowledge portal for contribution and the annotation of diverse metamaterial datasets using the Dataset Catalog (DCat) [1], Semantic Data Dictionaries [5], and data visualization tools, like Vega Lite [6], to provide a semantic resource for computational metamaterials research.

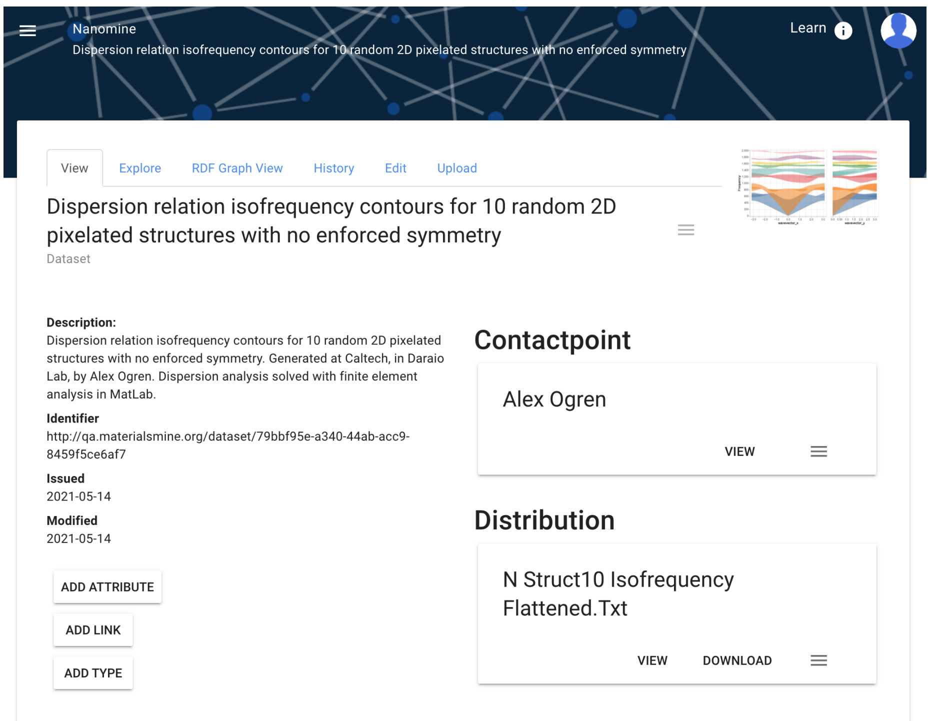

Fig. 1. All entities within the MaterialsMine knowledge graph have informational pages available to view. This will display relevant attributes (left) and links to other entities (right). Here we see a summary view of a metamaterial dataset for 2 dimensional structures.

2 Dataset Upload and Interpretation

The first part of the demo will showcase the ability to upload, annotate, and ingest experimental and simulation data into the MaterialsMine knowledge graph. Users are able to upload datasets and describe their metadata using the DCat vocabulary, and have it published as an information web page (see Figure 1), as well as 5 star linked data. The data files themselves can be referenced as downloadable resources using the link provided at "download". We will also demonstrate how users can link data files to Semantic Data Dictionaries to allow the knowledge graph to interpret those files as RDF graphs.

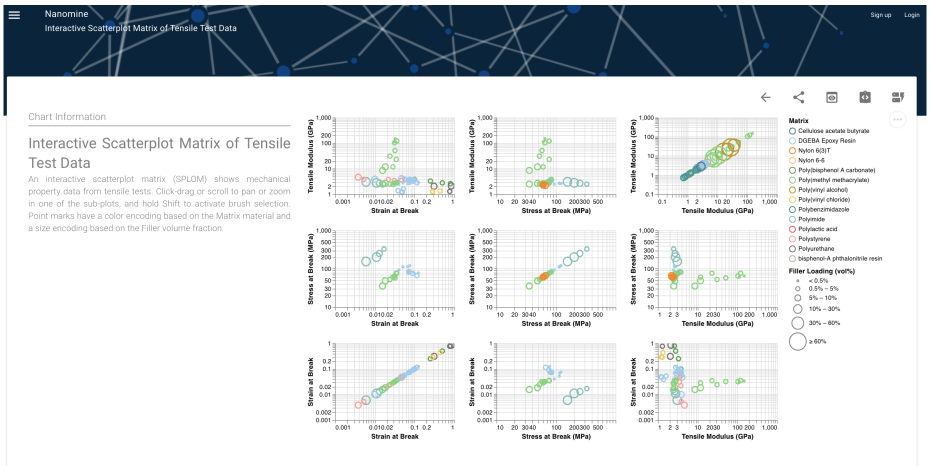

Fig. 2. An example interactive scatterplot matrix of nanomaterial tensile properties. Each visualization is reproduced from a query and visualization specification on-demand against the current knowledge graph.

3 Visualization

Most researchers, when visiting MaterialsMine, will be exploring existing data, and we will demonstrate how users can use the faceted browser, view, create, and share custom visualizations using Vega Lite, and explore existing queries using Data Voyager. Materials scientists with some training in SPARQL and Vega Lite are able to create custom visualizations (see Figure 2). Our demo will show how users can create and publish visualizations from SPARQL queries and uploaded data. Users without that experience can explore the data queried for existing charts to make their own visualizations using Data Voyager (see Figure 3). These capabilities, along with the faceted browser, create many opportunities for data exploration and analysis.

4 Availability

The MaterialsMine knowledge graph is published at http://materialsmine.org, using the Whyis knowledge graph framework. The MaterialsMine ontology is

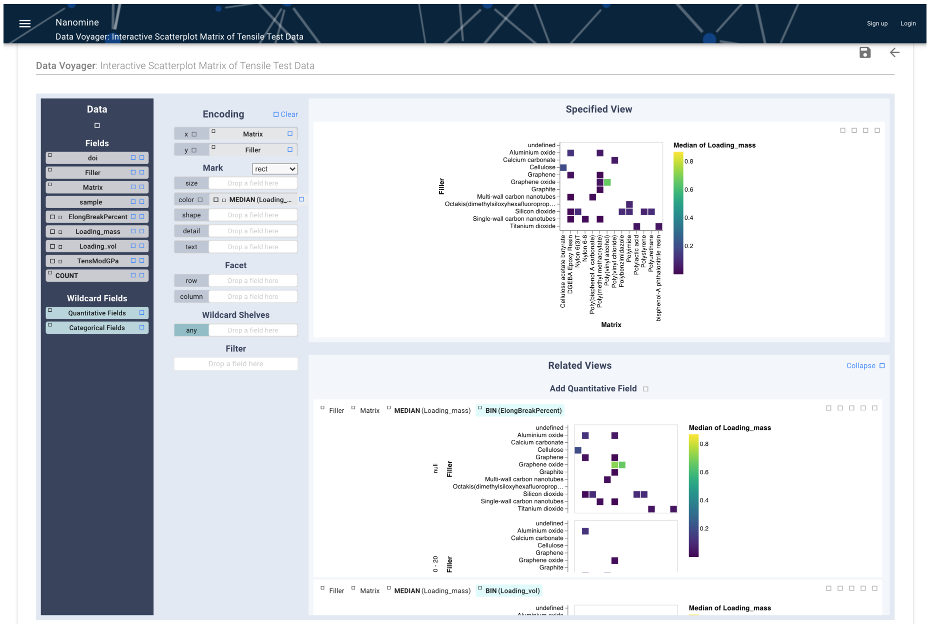

Fig. 3. Exploring the data stored in the interactive scatterplot matrix from Figure 2 using Data Voyager.

published as part of the knowledge graph and is available at http://materialsmine.org/ns. All entities in the graph, including datasets, are published as 5 star linked data aligned with SIO, Dublin Core Terms, and the W3C Provenance ontology, PROV-O. A read-only SPARQL API is available at https://materialsmine.org/wi/sparql, providing access to all material data and its provenance, as well as how it was transformed into RDF. Visualizations can be browsed at http://materialsmine.org/wi/gallery.

References

1. Consortium, W.W.W., et al.: Data catalog vocabulary (dcat) (2014)
2. Knittel, C., Kamien, R., Ghita, O., Evans, K., Dion, G.: Top down modeling of complex knit structures: Beyond jersey knits. In: APS March Meeting Abstracts. vol. 2019, pp. K63–005 (2019)
3. McCusker, J.P., Keshan, N., Rashid, S., Deagen, M., Brinson, C., McGuinness, D.L.: Nanomine: A knowledge graph for nanocomposite materials science. In: International Semantic Web Conference. pp. 144–159. Springer (2020)
4. McCusker, J., Rashid, S.M., Agu, N., Bennett, K.P., McGuinness, D.L.: The whyis knowledge graph framework in action. In: International Semantic Web Conference (P&D/Industry/BlueSky) (2018)
5. Rashid, S.M., McCusker, J.P., Pinheiro, P., Bax, M.P., Santos, H.O., Stingone, J.A., Das, A.K., McGuinness, D.L.: The semantic data dictionary–an approach for describing and annotating data. Data intelligence **2**(4), 443–486 (2020)

6. Satyanarayan, A., Moritz, D., Wongsuphasawat, K., Heer, J.: Vega-lite: A grammar of interactive graphics. IEEE transactions on visualization and computer graphics **23**(1), 341–350 (2016)
7. Zhang, B., Luo, Y., Liu, X., Barbastathis, G.: Macroscopic invisibility cloak for visible light. Physical Review Letters **106**(3), 033901 (2011)