Enhancing Materials Informatics through Ontologies and Functional Decomposition: Improving Data Traceability and Understanding

Hiori Kino
National Institute for Materials Science, 1-1 Namiki, Tsukuba, Ibaraki, Japan
Email: kino.hiori@nims.go.jp

ID: INV01
Invited Talk
Aug 23, 10.50 - 11.20

Materials informatics is an interdisciplinary field, and it is difficult to understand as a whole because there are specialized areas even within materials scientists, as well as data scientists working on databases, scientists working on machine learning methods. Because of the difficulty of understanding disciplines, ontologies formally define concepts and their relationships in order to facilitate data sharing and reuse. Even so, it is difficult to create an exhaustive lexical hierarchy of concepts in materials science. This approach can be used in a simple way. In addition, other methods developed by ontology specialists can be used to improve the explanatory power of the ontology.

In materials science, data are not immediately ready for analysis upon collection. They require data cleansing, which is purpose-dependent. Efficiently cleansing large datasets necessitates identifying data nodes that fail to be transformed appropriately. The NIMS Atomwork Advanced Database is optimized for browser display, not data analysis. Using this database as an example, I describe the data structure with ontologies via Protégé and transform it into a graph database. This method improves data traceability, clarity, and understanding.

It is difficult to describe the world using the relation-based description method in which ontology excels. For this reason, ontology experts have proposed a functional decomposition tree, which describes the world in a purpose-oriented and recursive manner, using natural language to describe transformations as the main subject. This method can be used to hierarchically describe tacit knowledge that is frequently omitted in expert explanations and to improve the understanding of the people involved. Even if you do not actually use this technique, simply being aware of it can improve explanatory performance. For example, a lexical hierarchy of material science concepts can be defined by a functional decomposition tree. Another example is to describe a program in terms of a functional decomposition tree.

Keywords: Materials Informatics, Ontologies, Functional Decomposition, Data Cleansing, NIMS Atomwork Advanced Database, Protégé, Data Traceability, Knowledge Hierarchy