Computers in Industry 73 (2015) 8–22

Contents lists available at ScienceDirect

Computers in Industry

journal homepage: www.elsevier.com/locate/compind

A survey on knowledge representation in materials science and engineering: An ontological perspective

Xiaoming Zhang [a,*], Chongchong Zhao [b], Xiang Wang [a]

[a] School of Information Science and Engineering, Hebei University of Science and Technology, 26 Yuxiang Street, Shijiazhuang 050018, Hebei, China
[b] School of Computer and Communication Engineering, University of Science and Technology Beijing, 30 Xueyuan Road, Beijing 100083, China

ARTICLE INFO

Article history:
Received 16 February 2015
Received in revised form 30 June 2015
Accepted 14 July 2015
Available online 25 August 2015

Keywords:
Materials knowledge
Ontology
Information integration
Linked data
Materials informatics

ABSTRACT

Informatics in materials science and engineering is playing more and more important role in contemporary society. Intelligent applications such as new materials discovery, materials selection, life prediction, and failure analysis heavily depend on diverse materials information and knowledge. However, most of the existing materials information scatters in different heterogeneous sources without proper interconnections. Therefore, the integration of materials information from heterogeneous sources and transforming them into materials knowledge are of great significance. For this reason, some materials ontologies have emerged, although they may have different motivations and application scenarios. In this paper, we survey the representation of materials knowledge in recent decade especially from an ontological perspective. For better investigation, the possible roles that materials ontology may play are defined in a hierarchical structure, and comparison criteria for the survey are given, which primarily focuses on ontological aspects and materials application aspects, respectively. Then, we review nine selected materials ontologies, and analyze the main features of them. Afterwards, the surveyed materials ontologies are further discussed and compared according to the comparison criteria and roles definition presented by this paper. Finally, the future challenges of materials ontology development and application are discussed.

Contents

* Corresponding author. Tel.: +86 13513386902.
E-mail addresses: zxmhebust@163.com (X. Zhang), zhaocc@ustb.edu.cn (C. Zhao), wangxiang@hebust.edu.cn (X. Wang).

http://dx.doi.org/10.1016/j.compind.2015.07.005

1. Introduction

Materials science and engineering is closely interrelated to our daily life as well as industrial production and manufacturing, so it is usually considered as one of the pillars of human civilization. Over the past decades, industrial community and research institutions have generated and accumulated immeasurable materials data which are sure to imply huge domain knowledge which is important to materials scientists. These materials knowledge makes it possible to develop intelligent applications such as new materials discovery, life prediction, failure analysis, fault diagnosis, and materials selection. For example, Granta CES selector [1], which is a software tool for materials experts to explore materials property data, provides a series of functions which can help users complete the tasks such as materials data comparison, materials selection, and materials development, in a very convenient and visualized way.

However, some reasons lead to the difficulties of utilizing the existing materials information. First, the relationships in materials information are complicated. Processing, structure, properties, and performance are usually considered as the four factors of MSE (Materials Science & Engineering), so materials information should not be isolated but interconnected. Second, since materials information may be application-specific, heterogeneous and dispersed, users can hardly find the required information even if they exist dispersively in the different data sources. Third, the classification methods of materials are diversified according to different point of view, but usually only one aspect is concerned in a specific materials data source. Lastly, materials knowledge hidden in the heterogeneous data source may not be discovered easily.

From above, it can be seen that it is hard to make full use of the diverse materials information if they are not linked with each other properly. That is to say, materials information should be integrated semantically, which calls for a unified representation strategy to make materials information work together. As we know, in relational data model [2,3], although we can define the relationship between two data tables by using foreign key, the semantics of this kind of links are implicit rather than explicit. Therefore, relational database is not enough for semantic integration of materials information. Fortunately, Semantic Web [4] technologies (e.g., ontology, logic rule) offer the ability to resolve the heterogeneity across different materials information sources. In recent decades, researchers in materials science as well as computer science have made a lot of efforts in this aspect, and semantic representation of materials information and knowledge becomes one of the key factors of materials informatics [5,6] which is about utilizing information technologies (e.g. data management, information integration, data mining, visualization) for the research in materials science. A task and working group of CODATA, named "Exchangeable Materials Data Representation to support Scientific Research and Education" [7,8], was approved in 2006, and one of its objectives is to promote open standardization to describe materials information. ELSSI–EMD [9] is a one-year CEN workshop (2009–2010), which aimed to develop schemas and ontologies based on materials testing standards. Materials Genome Initiative [10] of USA focuses on accelerate the process of discovering new materials based on materials data and tools.

Compared with Semantic Web application in other science and engineering area such as bioinformatics [11–13], biomedical informatics [14–16], chemical informatics [17–20], geographical informatics [21–23], and engineering informatics [24–26], materials informatics is still very young. The development and application of the semantic representation of materials knowledge is still on working. In biomedical field, there are lots of ontologies surveys [16,27–30] from different perspectives, but as for materials science, there are few surveys about materials ontology although some literatures [9,31] involved some related contents. Cheung, et al. [31] give a short review about some data integration methods and ontologies for materials science in its related work section. [9] introduces the existing schemas such as ISO 10303-235, MatML [32,33], and JRC MatDB [34,35], and discusses the merits and limitations of these schemas.

In this review, we investigate the representation methods of materials knowledge in recent years especially from an ontological perspective to show the state of the art as well as the future challenges in this area. The possible roles played by materials ontology are defined in a hierarchical structure, and meanwhile the comparison criteria for the survey are given, which primarily focuses on two views (i.e., ontological aspects and materials application aspects). Then, the most widely referred nine materials ontologies are reviewed and analyzed. Further discussion and comparison of these materials ontologies are done according to the comparison criteria and roles definition presented by this article. Finally, the future challenges of the representation of materials knowledge are presented.

The remainder of this paper is organized as follows. Section 2 outlines the related terminologies and conceptions. In Section 3, the roles of materials ontology are defined. Section 4 describes the