OOML-Based Ontologies and Its Services for Information Retrieval in UDMGrid*

Xixi Luo and Xiaowu Chen

The Key Laboratory of Virtual Reality Technology, Ministry of Education,
School of Computer Science and Engineering, Beihang University,
Beijing 100083, P.R. China
{luoxixi, chen}@vrlab.buaa.edu.cn

Abstract. In order to effectively integrate and share the enormous dispersed resources of various digital museums, University Digital Museum Grid (UDMGrid) has been developed to provide one-stop information services about kinds of digital specimens in the form of grid services. To eliminate the heterogeneity between the information resources, shared concepts for these digital museums are indispensable. This paper studies OOML-based ontologies and its services for information retrieval in UDMGrid, including the object oriented ontology construction and ontology mapping, in which a novel inheritance mechanism is proposed to eliminate logic confusion. On the basis of OOML-based ontologies, ontology services are developed to assist the information retrieval by transforming global concepts to local concepts.

1 Introduction

Eighteen featured university museums have been digitized mainly relating to Geology & Geography, Archaeology, Humanities & Civilization, and Aeronautics & Astronautics [1]. These digital museums play an important role in the fields of education, scientific research, as well as specimen collection, preservation, exhibition, and intercommunication. However, these digital museums dispersed on different nodes in CERNET (China Education and Research Network) [2] confront a problem that the multi-discipline resources at these digital museums are isolated and dispersed without sufficient interconnection. Hence it is necessary to propose a digital museum resources integration solution, through which these eighteen digital museums would be incorporated as a more comprehensive virtual one.

University Digital Museum Grid (UDMGrid) [3] [4] is proposed to using the grid technology to integrate and share the distributed digital museum information resources. From the user's perspective, UDMGrid should perform as a virtual digital museum, in which users can browser the digital specimen information in multiple manners on only one UDMGrid portal instead of eighteen separate homepages, with-

* This paper is supported by China Education and Research Grid (ChinaGrid)(CG2003-GA004 & CG004), National 863 Program (2004AA104280), Beijing Science & Technology Program (200411A), National Research in Advance Fund (51404040305HK01015).

J. Cao, W. Nejdl, and M. Xu (Eds.): APPT 2005, LNCS 3756, pp. 342 – 352, 2005.

out rushing about among these digital museums. However, the information resources of digital museums are constructed by different domain experts, who use special metadata to describe information, thus the heterogeneity among these metadata brings difficulties for users to locate, organize and integrate the information resources. Therefore, shared concepts for these digital museums are indispensable.

Ontology defines a set of representational terms that we call concepts, among which the interrelationship describes a target world [5]. In grid environment, Ontology becomes increasingly crucial to operations about the analysis and integration of information resources [6].Ontology is widely applied to information system. Knowledge Sifter is a scaleable agent-based system that supports access to heterogeneous information sources such as the Web, open-source repositories, XML-databases and the emerging Semantic Web, in which the concept of ontology is central to this approach. By the Ontology Agent, user can pose queries to those data sources without needing to know the location of the supporting data, nor how the ontological concepts are materialized through the integration and ranking process [7]. Infosleuth is a retrieval agent system that provides access to information form multiple domains, regardless of its heterogeneity or distribution, in which domain ontologies and an evolutionary model of the use's interests are some of the basic concepts used by the system to help users identify and retrieve relevant domain information [8]. Meanwhile, the application of ontology in grid is emerging. Earth System Grid, a project of the U.S. Department of Energy Scientific Discovery through Advanced Computing (SciDAC) program, used and modified ontological concepts for its domain area to provide a basis for classifying and retrieving data files, collections and information about the files and collections based on content for use in a grid context [6]. Furthermore, some ontology development methodology comes out. ROD, a rapid ontology development methodology that can be used to build ontology for underdeveloped domains, make the development process more efficient so that domain concepts and relations can be automatically discovered from large-scale semi-structured and/or unstructured textual resources[4]. Additionally the ontology modeling is up-and-coming. OIL [9], DAML [10], SHOE [11], RDF(S) [12] and OOML (improved Object Oriented Markup Language) [13] are introduced as ontology modeling. Compare to former four kinds of makeup language, OOML are suitable for the UDMGrid ontology modeling, since the better support of generalization and specialization goes well with the ontology internal structure of UDMGrid.

This paper studies the OOML-Based Ontologies and its services for information retrieval in UDMGrid, through which a unified view of available resources is provided to users. In this paper, OOML is adopted for object oriented ontology modeling, in which a novel inheritance mechanism is proposed to eliminate logic confusion. Based on the mechanisms of ontology mapping in UDMGrid, ontology services are developed to assist the information retrieval.

The remainder of this paper is organized as follows. Section 2 presents the modules concerning ontology in UDMGrid. Section 3 elaborates the OOML-based ontologies in UDMGrid, including the hierarchy architecture and the mapping strategy. On the basis of OOML-based ontologies, the ontology services are introduced in Section 4, and Section 5 briefly issues the application workflow. Section 6 ends this paper with conclusions and future work.