Practice-Based Ontologies: A New Approach to Address the Challenges of Ontology and Knowledge Representation in History and Archaeology

Emad Khazraee and Michael Khoo

Drexel University, College of Information Science and Technology, 3141 Chestnut St., Philadelphia, PA, 19104, USA
{Emad,khoo}@drexel.edu

Abstract. Data production in history and archaeology far outpaces data processing. In order to apply computers to this problem, historical data must be converted to machine-readable forms. This process is easy for domains of knowledge that have explicit terminology, but history and archaeology lack these characteristics. This study therefore proposes a phenomenological approach to requirements gathering for knowledge representation and ontology systems for historians and archaeologists. The approach utilizes qualitative and ethnographic research methods to gather data on practitioners’ reasoning and knowledge practices. The design requirements for ontology design can be extracted from the ‘thick description’ produced by this process, and used to build ‘practice-based ontologies.’ This paper presents the theoretical framework and early outcomes of ethnographic research with archaeologists in practice at the University of Pennsylvania.

Keywords: Ontology, knowledge representation, practice-based ontology, ethnographic method, archaeology.

1 Introduction

Historical and archaeological data are very diverse. They are dispersed in different institutions, and across different countries. The pace of data production in this field is far higher than the amount of data being processed. This situation leaves large amounts of data to be processed, and many research questions unsolved. Increasing computational power can help to solve more complicated questions, but a precondition to benefit from computers is that data must be converted and formalized into machine-readable formats.

Knowledge engineering is “the application of logic and ontology to the task of building computable models” that “analyzes knowledge about some subject and transform it to a computable form for some purpose” [1]. It includes the study of a domain and the mechanics and dynamics of knowledge in that domain, to ascertain whether it is possible to achieve computable forms for that purpose. Davis et al. [18] summarized following five principles for knowledge representation. A knowledge representation: is a surrogate, is a set of ontological commitments, is a fragment
theory of intelligent reasoning, is a medium for efficient computation, and is a medium of human expression. These principles are easier to achieve in domains of knowledge that have explicit terminology, well-formed definitions, generic forms of analyses, and techniques for verification. However, history as a domain to some extent lacks all the above characteristics [2-3]. Uncertainty, contradictory data, and different analytical perspectives are intrinsic characteristics of historical investigation. These make the historical reasoning process complex: the problem with history is that what we know is vague. Domains such as history and archaeology have therefore not been addressed very well in knowledge representation.

This paper introduces new directions for the study and design of knowledge representation systems and ontologies for history and archaeology. First, some existing systems, approaches and challenges for history and archaeology knowledge will be discussed. We then describe an opportunity for addressing some of these challenges through the ethnographic study of the communities of practice amongst archaeologists. We call this approach ‘practice-based ontology design.’ The findings of a pilot empirical study will be presented.

2  Ontologies for History and Archaeology

This paper considers more recent approaches to archaeology known as post-processual and interpretative archaeology. Processual archaeology was inspired by “hypothetic-deductive positivism derived from Hemple (deducing statements from general theories and test them against observable data)” (Renfrew & Bahn, 2005, p. 208). In contrast, post-processual archaeology emphasizes the interpretation instead of explanation. In this paradigm the interpretation process is done considering hermeneutic cycle, a back and forth process between interpretation and data, between whole and part. Fitting is used instead of hypothesis testing in this paradigm. The main goal of this paper is to understand the requirements for ontologies to support the generation of cohesive interpretive narratives from archaeological data processing.

A number of ontologies already exist to facilitate information sharing and exchange, including the CIDOC Conceptual Reference Model (CRM) for cultural heritage, developed by the International Council of Museums (ICOM) and adopted as ISO standard in 2006 [5-6]. However, these ontologies are not enough to accelerate historical data processing. A more ambitious vision for ontologies in history and archaeology includes knowledge acquisition, reasoning, hypothesis generation and verification, and narrative storytelling. It is therefore necessary to study how to employ formal ontologies for these goals. First however we need to understand the characteristics of ontologies for history and archaeology, and how these characteristics differ from those of scientific domains.

2.1  Complexities and Obstacles of Knowledge Representation in History and Archaeology

The nature of historical investigation is different from that of scientific investigation. The works of historians is more like the work of detectives, findings and assembling clues from the available data and proposing a possible reconstruction of the past [5].