Information and Organization 19 (2009) 199–217

Contents lists available at ScienceDirect

Information and Organization

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

Between meaning and machine: Learning to represent the knowledge of communities

David Ribes [a,*], Geoffrey C. Bowker [b]

[a] Georgetown University, Communication, Culture and Technology, 3520 Prospect St. NW, Suite 311, Washington, DC 20057, United States
[b] Santa Clara University, Science Technology and Society, 500 El Camino Real, Santa Clara, CA 95053, United States

ARTICLE INFO

Article history:
Received 4 June 2008
Received in revised form 25 March 2009
Accepted 20 April 2009

Keywords:
Ontology development
Routine
Semantic interoperability
Sociology of knowledge representation
Ethnography

ABSTRACT

Representing knowledge in codified forms is transformative of ones orientation to that knowledge. We trace the emergence of a routine for knowledge acquisition and its consequences for participants. Over time, participants in the earth science project GEON, first learned about ontologies and then learned how to create them. We identify three steps in the routine: understanding the problematic of interoperability; learning the practice of knowledge acquisition; and engaging the broader community. As participants traversed the routine they came to articulate, and then represent, the knowledge of their communities. In a process we call reapprehension, traversing the routine also transformed participants' orientation towards their data, knowledge and community, making them more keenly aware of the informational aspects of their fields.

1. Introduction

Ontologies are an information technology for representing specialized knowledge in order to facilitate communication across disciplines, share data or enable collaboration. In a nutshell, they describe the sets of entities that make up the world-in-a-computer, and circumscribe the sets of relationships they can have with each other. They are a complex and ambitious technical approach to address the problem of diverse languages, heterogeneous categorizations and varied methods for organizing information. In the wake of ontologies the information of a domain is substantially reorganized, facilitating data exchange and reuse. These are the *goals* for ontologies. Their *development* is a practical and organizational achievement, and the topic of this paper. We focus on the practical processes surrounding

* Corresponding author. Tel.: +1 202 687 4831.
E-mail addresses: dr273@georgetown.edu (D. Ribes), gbowker@scu.edu (G.C. Bowker).

1471-7727/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.infoandorg.2009.04.001

the design and deployment of ontologies within the infrastructure project GEON (the geosciences network) and trace the emergence of an organizational *routine* for their production. This routinization mirrored the learning trajectory of participants as they came to understand what is *at stake* in ontologies, that is, that they were representing the knowledge of their communities. As they traversed the routine, participants' experience of expert knowledge and of their communities were transformed in a process we call *reapprehension:* an increased orientation to the informational organization of ones field e.g., data, databases, encoded knowledge and their capacity to "flow" (or interoperate) across technological, disciplinary and institutional divides.

Organizational routines are "repeated patterns of interdependent actions, performed by multiple actors," (Feldman, 2000; Feldman & Pentland, 2003). They serve as a resource, a malleable and locally adapted recipe or template for how to go about a task. While a routine must always be practically enacted, having no existence outside its performance, it also becomes embedded in the configuration of material resources that enable practical work (Jordan & Lynch, 1998). We will see that the activities we trace and call the *acquisition routine* rendered the complex and uncertain activities of knowledge representation into an outline of "steps," reducing the work of reinvention on each new occasion of ontology building. However, traversing the routine also changed the experience participants had of their data archive, knowledge and community. In particular we focus on the emergence of a structured concern for ensuring the representativeness of ontologies: the practical methods for creating representations which *stood in for* and *were used by* a larger knowledge community.

Ontologies are intended to serve a community, making accessible data and resources for its members; as such they are a form of *infrastructure* (Star & Ruhleder, 1994). Participants in GEON quickly realized that they comprised a small subsection of the geoscience community (i.e., scores of participants in a community often cited to be in the thousands). Without the work of making ontologies representative of their domain community – of generating venues for feedback and for participation – their ontologies would be open to contestation, or, more likely, be ignored and remain unused. However, modeling ontologies involved articulating knowledge in ways that appeared alien to that domain community. For ontologies to appear representative, the community itself would have to learn the goals and language of knowledge modeling.

The phenomenologically transformative consequences of learning and traversing routines are almost completely undiscussed in the literature. In our case, participants' orientation to information in their discipline was changed by traversing the routine. We name this reapprehension, and emphasize the practice and material tools that accompany the reworking of, for instance, knowledge in informational terms. Participants came to learn: (i) the purposes and goals of ontology, what we call the problematic of interoperability; (ii) how to articulate their knowledge in forms amenable to formal representation, and (iii) how a broader community's interests are at stake in this process, and what activities would be necessary to engage and enrol that community in the use of ontologies. In order to do so participants had to rearticulate their knowledge in forms amenable to formal modeling, and also encourage their colleagues in the ontologies' use, maintenance and upkeep. A keener awareness of the informational aspects of their fields changed the orientation of participants to their own data and knowledge; it also entailed redirecting more time and resources to their integration and maintenance.

Following a discussion of case, method and an outline of knowledge capture we trace each of the three steps of the routine. In Section 6 we return to how, by traversing the routine, "knowledge" and "community" took on new meaning, as they were rearticulated in the language of logic and information as predicates and users,[1] respectively.

1.1. Case and method

GEON, the GEOscience Network, is a cyberinfrastructure project (Atkins, 2003) which sought to produce a repertoire of high-end information technologies for the broader earth sciences:

[1] There is a close relationship between acquisition and user studies or requirements analysis. In both cases people are recast as users of future systems, the object of studies that make them known, so as to inform a process of technology design (Mackay, Carne, Benyon-Davies, & Tudhope, 2000; Woolgar, 1991). This topic is analyzed more extensively by Ribes and Finholt (2008) in which the authors explore the simultaneous constitution and knowing of a user community.