Practical Ontologies for Information Professionals by David Stuart (review)

David B. Levy
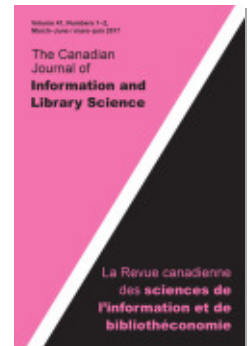
Reviews
Comptes-rendus de lecture

Stuart, David. *Practical Ontologies for Information Professionals*. London: Facet Publishing, 2016, 184 pages. £59.95.

This book is published at a pivotal sea change in the history of library information and the history of ontologies. Todays' information landscape is experiencing a flood of streamed data from new online social behaviours, leading to the existential crisis quoted by Stuart at the beginning of his book: "Drowning in information, but being starved for knowledge" (p. 1). Stuart demonstrates the importance of fighting information overload by the development of robust and widely used ontologies. Stuart's aims in this book are threefold: first, to demonstrate the importance of ontologies for knowledge discovery to information professionals; second, to describe the important contributions information professionals can make to ontology development; and, third, to provide a practical introduction to the development of ontologies for information professionals. The book also introduces information professionals to basic concepts and tools they can use to build new ontologies and question existing ontologies. Stuart acknowledges that no book could provide an exhaustive introduction to the practicalities of ontology use and development. However, he does a good job on focusing on underlying theory and principles related to similar technologies that can be applied in the field of ontologies. In addition, Stuart acknowledges that, while library professionals are more likely to use ontologies than develop them, it is important that they be aware of some of the decisions that go into the development process.

Stuart pulls off the Herculean task of defining the various equivocal uses of the term ontology. He notes: "Ontologies are one way of helping to tame some of the problems identified above, providing a structure for this information in such a manner that it can be read automatically and unambiguously and shared more widely" (p. 4). Ontologies, or formal representations of knowledge with rich semantic relationships, emerged as a way of capturing knowledge and codifying it in a highly structured manner as data. As a result, they are tools that researchers can use to extract maximum value from data. In addition, they are one of several different knowledge organization systems that have been developed within the information profession to improve information discovery. Ontologies are thus related, but different from, knowledge organization systems such as controlled vocabularies, synonym rings, thesauruses, taxonomies, facet analysis, indexing, and hierarchical term trees.

The semantic Web is a key factor in the increased interest in ontologies. Ontologies that are publicly available and widely used as part of the semantic

Web are the primary focus of much of the book. The potential of ontologies is, in part, based around synthesizing and bringing together distributed resources, and the Web provides the largest potential number of contributors to such a system. As a result, the need to know about ontologies will increase in the future. Ontologies have the potential to make information retrieval more accurate and efficient as well as to enhance the ability of finding undiscovered public knowledge. The four principle purposes for taxonomies equally apply to ontologies: (1) indexing support; (2) retrieval support; (3) organization; and (4) navigation support. The practical implications are many. An ontology enables an indexer to think more broadly about the terms that are being applied, with a wider range of associated applicable terms. Ontologies can improve information and knowledge retrieval and help with the mining of undiscovered public knowledge in an increasingly automated fashion.

The book is structured in six chapters that lead from the introduction of the semantic Web and some existing ontologies, through the adoption, building, and interrogation of ontologies, to a vision of their increasing importance. Chapter 2 looks at ontologies in relation to the semantic Web. It provides an introduction to the semantic Web and how ontologies have been adopted by libraries, cultural heritage institutions, and commercial businesses. Chapter 3 considers some of the main existing ontologies, including those ontologies used for representing ontologies, those widely adopted by libraries, and those widely used on the Web. Stuart acknowledges that many ontologies appear to compete for attention on the semantic Web. Thus, most ontologies may be destined to expire before they are adopted, and even those that are widely adopted will not necessarily be used immediately. With that caveat, this chapter introduces some of the most important ontologies for the semantic Web. Finally, Stuart clarifies the interrelated nature of many ontologies, highlighting that application profiles often reuse elements.

Chapter 4 focuses on adopting ontologies. It shows the reuse of existing ontologies for both the integration of data across different systems and to avoid duplication of work. In addition, some of the tools that are available for identifying existing ontologies, examining how ontologies can be combined to make application profiles, and deciding which criteria to consider when selecting ontologies are also explored. Stuart demonstrates that that there is a need for higher-quality ontology libraries and search engines if we are to move beyond the stage where many of the ontologies fall into disuse before being widely adopted.

Chapter 5 addresses the development of ontologies for particular applications. This chapter provides both a method for building an ontology and an overview of some of the tools that are available. Stuart also leads the reader through the development of a simple ontology with Protégé, a free and popular ontology development software. In this chapter, Stuart considers some of Dean Allemang and James Hendler's errors when modelling for the semantic Web, such as rampant classism, creeping conceptualization, exclusivism, and objectification.[1]

Chapter 6 interrogates ontologies for the often helpful data they contain. It provides an overview of tools available for questioning semantic Web ontologies

through Simple Protocol and RDF Query language (SPARQL) and web crawlers to gain knowledge. This chapter shows a wide range of tools available for exploring ontologies and the semantic Web, although many of them continue to be aimed at computer scientists or professional ontologists. The type of tool to be used is often dependent on the size of the ontology, its complexity, and the documentation as well as the number of classes and properties.

Chapter 7 looks to the future of ontologies and the information professionals' development and use of them. It forecasts that the development of ontologies is likely to be integrated with other technologies, such as natural language processing and the potential to crowd-source workflows. Stuart predicts that the development of niche subject-specific ontologies in growing in the long term. Today, most search engines primarily return pages rather than answers, and many of those pages are false drops. Something as simple as generating results to compare prices can have blind spots. Stuart notes that the way we represent, publish, and retrieve data will continue to change as new technologies emerge. Stuart cautions that ontologies will not replace older forms of controlled vocabulary, but neither will ontologies be replaced by natural language processing. Ontologies are one of a wide variety of tools that promise to enable new insights from large data sets.

Stuart finishes his book with a note of caution and concern. He acknowledges that there are many challenges posed by semantically rich data sets, such as our notions of privacy in a world connected by ontologies. Such challenges make it essential that library and information professionals be informed regarding ontologies. This book is recommended for computer scientists, information professionals, and epistemologists concerned with the nature of knowledge during the technological revolutions of the twenty-first century.

David B. Levy, Touro College

Note

1   Dean Allemang and James Hendler, *Semantic Web for the Working Ontologist*, 2nd ed. (Boston: Morhan Kaiufmann, 2011).