IDEA GROUP PUBLISHING **ITB11172**

701 E. Chocolate Avenue, Suite 200, Hershey PA 17033-1240, USA
Tel: 717/533-8845; Fax 717/533-8661; URL-http://www.idea-group.com

This chapter appears in the book, *Business Systems Analysis with Ontologies*, edited by Peter Green and Michael Rosemann. © 2005, Idea Group Inc.

Chapter VIII

Using a Common-Sense Realistic Ontology: Making Data Models Better Map the World

Ed Kazmierczak, The University of Melbourne, Australia

Simon Milton, The University of Melbourne, Australia

Abstract

This chapter examines the following question: "How well do data models map the world?" Data modelling languages are used in today's information systems engineering environments to model reality. Many have a degree of hype surrounding their quality and applicability with narrow and specific justification often given in support of one over another. We want to more deeply understand the fundamental nature of data modelling languages. We thus propose a theory, based on ontology, that should allow us to understand, compare, evaluate, and strengthen data modelling languages. We then introduce Chisholm's ontology and apply methods to analyse some data modelling languages using it. We find a good degree of overlap between all of the data modelling languages analysed and the core concepts of Chisholm's ontology, and conclude that the data modelling

languages investigated reflect an ontology of commonsense-realism. Critical common-sense realism more generally due to its perspectival nature and its implicit recognition of institutional and social reality has the potential to dramatically improve our ability to better map the world.

Introduction

Data models have been used in information engineering environments for many decades for the precise purpose of building representations of reality. Data models are used in organizations to represent reality at three levels. First they are used *to establish the highest level of description of an organisation's reality* to guide strategic information systems development and management and for high level data management. The model is used to drive information systems management and development and for the implementation or management of databases. Second, they are used *to construct a description of the reality surrounding a proposed information system*. The description is used in systems analysis and design. This is often called "conceptual modelling" although the name does not describe well the purpose of this activity. This facilitates the accurate and timely implementation of a system by helping establish relevant shared understandings of reality and in implementing some specific aspects of the system in technology such as databases. An increased degree of detail is required compared with corporate data modelling. Significant attributes of things found in the reality are required together with relationships between entities. Finally, they are used *to model parts of an organisation's reality leading to implementation in an operational database* into which facts about that reality are stored. Such databases may serve several information systems within an organization. This description is the most detailed but assumes only enough detail for all applications relying on it to function.

To date, there have been many different data modelling languages proposed with the most popular being the entity-relationship model (Chen, 1976) but also including the functional data model (Kerschberg & Pacheco, 1976; Shipman, 1981), the semantic data model (Hammer & McLeod, 1981), NIAM (Nijssen & Halpin, 1989), and object modelling technique (Blaha & Premerlani, 1998) that later became the basis for the unified modelling language (UML). Each new modelling language has often been accompanied with claims of its superiority and at times hype when compared with the others. There has been little beyond opinion to substantiate such claims and yet all notations purport to do similar things. We have two research questions: