Chapter X

Toward an Extended Framework for Human Factors Research on Data Modeling

Heikki Topi, Bentley College, USA
V. Ramesh, Indiana University, USA

ABSTRACT
This study reviews and synthesizes over 15 years of research on human factors issues in conceptual data modeling. In addition to analyzing the variables used in earlier studies and summarizing the results of this stream of research, we propose a new framework to help with future efforts in this area. We also identify several key areas for future research and highlight the importance of building a strong theoretical foundation and using it to guide future empirical studies. It is our hope that this chapter allows both scholars and practitioners to utilize the results of existing research better and encourages continued work on conceptual data modeling.

INTRODUCTION
Conceptual data modeling continues to be an integral part of the foundation on which information systems are built. Depending on the development methodologies that are used for a particular project, the terms and methods used for conceptual data modeling vary, but in practice, a clear majority of methodologies used for systems development include a set of tools and methods for modeling data at the conceptual level. Therefore, it is not surprising that research in IS and its reference disciplines has shown a significant interest in various aspects of data modeling for the past 20 years. The focus of this chapter is on research that
examines the *usability* of various conceptual data modeling approaches, that is, research that investigates human factors issues in conceptual data modeling. We review and analyze this literature and suggest several new directions for further research.

**BACKGROUND**

The concept of *data modeling* has been used with a variety of different meanings within various areas of study and practice. However, within the organizational context the core idea underlying all the definitions is the same: A data model is used for describing entities\(^1\) and their relationships within a real-world domain. For example, McFadden, Hoffner, and Prescott (1999) define a data model as “an abstract representation of the data about entities, events, activities, and their associations within an organization”. A data model is an abstraction and a simplification of the domain it describes and thus, it always represents a limited part of reality.

The main focus of this chapter, *conceptual data modeling*, requires further clarification. Based on the ANSI/SPARC definition, a conceptual data model is any model that is independent of the underlying hardware and software. This means that using this definition, models created using formalisms ranging from the relational model to the semantically rich variants (Teorey, Yang & Fry, 1986) of Entity-Relationship modeling (Chen, 1976; Hull & King, 1987) can be considered to be at the conceptual level. A more restrictive definition of a conceptual model can be found in Batra and Davis (1992). They define a conceptual model as one that is capable of capturing the structure of the database along with the semantic constraints into a model that is easy to understand, does not contain implementation details, and can be used to communicate with users. A key criterion in the above definition is *the independence of modeling from the implementation technology*. This means that in order to be categorized as a conceptual model the representation must not be dependent on the characteristics of the database technologies available (e.g., relational, object-oriented, object-relational, network, or hierarchical).

We believe that both of the definitions presented above are, however, somewhat misleading because a true conceptual data model should capture the essential data characteristics of the *domain of interest*, and not necessarily the structure of the database. Thus, we define a conceptual data model as a set of constructs that can be used to create an abstraction of reality, that is, a representation capable of capturing the data-oriented (as opposed to process-oriented) aspects of a domain of interest in a manner that is unambiguous and easy to understand for analysts, designers, and users alike. Note that this definition does not have any references to a database structure. This is because we believe that not everything captured in a representation created using a conceptual data model will (or needs to) be reflected in a database or the eventual system being developed.

Based on the above definition of conceptual data modeling, one can synthesize at least five different uses for conceptual data models (Batra, Hoffer & Bostrom, 1990; Cambell, 1992; Juhn & Naumann, 1985; Kung & Solvberg, 1986): 1) a communication tool between analysts and users for the discovery (elicitation and representation) and validation stages of the systems analysis process, 2) a mechanism that helps analysts understand the domain of interest, 3) a formal conceptual foundation for organizational information systems at various levels (a common accepted model of reality and a communication tool between IS professionals, e.g., analysts and developers), 4) a foundation for applications developed by...
Applying UML and XML for Designing and Interchanging Information for Data Warehouses and OLAP Applications
www.igi-global.com/article/applying-uml-xml-designing-interchanging/3305?camid=4v1a

Towards a Comprehensive Concurrency Control Mechanism for Object-Oriented Databases
www.igi-global.com/article/towards-comprehensive-concurrency-control-mechanism/51156?camid=4v1a