Knowledge-Based Systems for Data Modelling

Sabrina Šuman, Business Department, Polytechnic of Rijeka, Rijeka, Croatia
Alen Jakupović, Business Department, Polytechnic of Rijeka, Rijeka, Croatia
Francesca Gržinić Kuljanac, City of Opatija, Opatija, Croatia

ABSTRACT

Data modelling is a complex process that depends on the knowledge and experience of the designers who carry it out. The quality of created models has a significant impact on the quality of successive phases of information systems development. This paper, in short, reviews the data modelling process, the entity-relationship method (ERM) and actors in the data modelling process. Further, in more detail it presents systems, methods, and tools for the data modelling process and identifies problems that occur during the development phase of an information system. These problems also represent the authors’ motivation for conducting research that aims to develop a knowledge-based system (KBS) in order to support the data modelling process by applying formal language theory (particularly translation) during the process of conceptual modelling. The paper describes the main identified characteristics of the authors’ new KB system that are derived from the analysis of existing systems, methods, and tools for the data modelling process. This represents the focus of the research.

KEYWORDS

Data Modelling, Entity Relationship Method, Formal Language, Knowledge-based System, Methods and Tools for the Data Modelling Process

1. INTRODUCTION

The information systems development process consists of the following phases: strategic planning, analysis, design, implementation, maintenance, and evaluation. The authors’ research interest is in the data modelling process that occurs in the design phase of information systems development.

As part of the information system design and implementation, the database structure has been developed. It serves to satisfy the information needs of information system users. This includes conceptual and logical database design (based on collected and analyzed requirements), physical database design and database implementation (Elmasri & Navathe, 2011). Conceptual and logical database design is in the focus of the author’s research.

The aim of this paper is to present the modelling process, the entity-relationship method (ERM), actors in the data modelling process and systems, methods and tools for the data modelling process.

Conceptual database design (conceptual data modelling) is an activity that is performed in the early stages of information systems development. Based on the user requirements, the process of the conceptual design identifies the fundamental concepts and relationships of observed reality and displays them in the conceptual schema form (Batini, Ceri & Navathe, 1992).

One of the main conceptual design problems is the creation of a complete, easy-to-use, understandable and semantically correct conceptual schema. Applying different methods in the conceptual design process attenuates this problem.
Each modelling method uses a small set of constructs (method vocabulary) and syntax rules (Siau, Wand & Benbasat, 1996). Constructs are graphical or non-graphical expressed concepts whose purpose is to organize and present the knowledge of the interest domain. For example, in the entity-relationship method (the most used method in conceptual design), the main constructs are entities, relationships, and attributes (Siau & Rossi, 2011; Shoval & Shiran, 1997).

The final aim of the authors’ main research is the development of a KB system for the data modelling process using finite state transducers from formal language theory (i.e. translation). This paper shows the results from related work analysis: presents, in short, the data modelling process, the entity relationship method (ERM) and the actors in data modelling process and, in more detail, systems, methods and tools for the data modelling process.

Related work analysis is based on papers that have been found in the databases EBSCO, ScienceDirect, IEEE, Google Scholar and on the scientific, social network ResearchGate. The expressions “knowledge-based systems for data modelling”, “formalisation of data modelling”, and “entity-relationship method” were used as search strings. The search results were further filtered by relevance to the authors’ new KB system. The relevance criteria included: user interaction methods, reasoning based on past cases, application of formal language theory (grammar, syntax analysis, parser), and rules for translation into entity-relationship model constructs. Finally, the search resulted in a total of 12 systems that support data modelling. Some of them have the characteristics of KB systems, while others are at the level of data modelling tools. The search also resulted in a number of methods that facilitate the creation of data models.

Including the introduction, this paper consists of five sections. The second section describes the research motivation. The third section shows the data modelling process and entity-relationship method. The fourth section describes the key actors in the data modelling process – potential beneficiaries of the authors’ new KB systems for data modelling. The fifth section shows some of the existing systems for data modelling support and some of the methods that facilitate the modelling process. The paper ends with a conclusion that specifies further research directions.

2. RESEARCH MOTIVATION

The data modelling research motivation has derived from a number of problems arising in the information systems development. Some of these problems in the analysis and design phase are:

- complexity of the process model to data model translation
- lack of data modelling knowledge
- non-existent or insufficient data modelling knowledge formalisation required in computer- systems
- error occurrence in data modelling
- low efficiency of the process model to the data model translation (time, money, and other resources)
- possible low effectiveness of the data model in the case of business process changes

The initial result of the data modelling process is a conceptual data model (created through conceptual modelling). Conceptual modelling focuses on the conceptual aspects of a domain of interest and, unlike database modelling, excludes considerations relating to the database design and implementation. One of the main research topics in the domain of data modelling is related to the evaluation of conceptual modelling grammar, such as entity-relationship or business process modelling grammar (Siau, & Rossi, 2011; Burton-Jones, Wand & Weber, 2009). It is considered that increased research focus on this type of grammar could result in more quality conceptual modelling scripts. In this case, a script is a text-expressed conceptual model that is generated by the language grammar (e.g. text-expressed ER diagrams generated by ER grammar). The terms appearing in this research
ERP and Beyond
www.igi-global.com/chapter/erp-beyond/48653?camid=4v1a

Security of Symbian Based Mobile Devices
Göran Pulkkis (2007). Advances in Enterprise Information Technology Security (pp. 21-74).
www.igi-global.com/chapter/security-symbian-based-mobile-devices/4790?camid=4v1a