

## Chapter 7.25

# Improving Storage Concepts for Semantic Models and Ontologies

**Edgar R. Weippl**

*Vienna University of Technology, Austria*

**Markus D. Klemen**

*Vienna University of Technology, Austria*

**Stefan Raffeiner**

*Vienna University of Technology, Austria*

### **ABSTRACT**

*Ontologies are more commonly used today but still little consideration is given of how to efficiently store them. The proposed approach is built on reliable and efficient relational database management systems (RDBMS). It can be easily implemented for other systems and due to its vendor independence existing data can be migrated from one RDBMS to another relatively easy.*

### **INTRODUCTION**

During the last couple of years ontologies moved into the center of interest in mainstream computer

science research. With the Internet becoming a truly global information resource, the effort required to find the right information increased, even though the quality of search engines improved considerably.

The next big step is anticipated to be the integration of semantic information of electronically available resources which will allow searches to obtain much better results. The process of building the required ontologies can either be top-down or bottom-up.

The RDF-based approach, favored by Tim Berners-Lee, strives to semantically enrich each Web page and build ontologies by integrating all the semantic information. Topic Maps, in contrast, are usually regarded as top-down approach where

occurrences are linked to topics once the Topic Map exists.

A prerequisite to building large ontologies is an efficient way of storing the required data. Today, it is generally agreed that ontologies evolve over time and require maintenance. Thus both retrieval and updates need to be handled efficiently by the storage system.

In this chapter we present an improved database schema to store ontologies. More specifically, our contribution is to:

- Propose an intuitive and efficient way of storing arbitrary relationships (Section 2.1)
- Show that our database schema is well suited to store both RDF and Topic Maps (Section 2.2)
- Explain why it is more efficient by comparing it to other approaches (Section 3)

## **LINK-BASED SCHEMA**

In this section we first explain the general idea of the improved database schema and provide an example of how concepts and relationships between them are stored. We then show how both RDF and Topic Maps can be stored, too.

### **Database Schema Based on a Link Table**

The idea, first described in (Weippl et al 2005), is based on an architecture that uses relational database. Tables are not linked to others directly with foreign keys or by using  $n : m$  intermediary tables but via a single, generic association table referred to as the link table.

In the classical schema, adding an  $n : m$  relationship between two tables requires creating a new intermediate table to resolve the  $n : m$  relationship into a  $1 : n$  and a  $1 : m$  relationship. Our approach is to merge these intermediate tables

into one link table which stores all relationships centrally.

The advantages of our approach are:

1. In contrast to classic E-R approaches, any relationship can be added without schema modifications. This allows to easily perform operations within transactions.
2. Tables and indices can be clustered to improve the speed of join operations with the central link table. In the classical model many  $n : m$  relationships exist, therefore, cluster optimizations are far more difficult and less efficient.
3. Our approach allows retrieving relationships from the link table without accessing the data dictionary. Since the data dictionary is vendor specific, the classical approach requires modifying the application for each database system.
4. If  $n$  entities exist and  $n : m$  relationships are to be established between all entities, the number of additional tables is  $O(n^2)$ , whereas our approach is  $O(1)$ . Of course this applies only to new relationships, not new tables.

Detailed explanations on the advantages can be found in (Weippl et al 2005).

Figure 1 shows a simple database schema. Table `ref1` contains the SQL statements of the following example. A new file type Document (.doc) is created with OpenOffice. An optional description is added and a relationship between the two topics is established (steps 1–4). In the same way occurrences can be linked to topics.

Steps 5–10 in Table 1 show how reification (Figure 2) can be implemented using our schema.

Our concept differs from other approaches (Section 3) by using separate tables to store different types of entities but one central link table for all relationships. The data-centric approach, which we also refer to as the 'classical' way, uses

9 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

[www.igi-global.com/chapter/improving-storage-concepts-semantic-models/8041](http://www.igi-global.com/chapter/improving-storage-concepts-semantic-models/8041)

## Related Content

---

### Measuring the Determining Factors of Financial Development of Commercial Banks in Selected SAARC Countries

Arodh Lal Karn, Girish Santosh Bagale, Bhavana Raj Kondamudi., Deepesh Kumar Srivastava, Ravi Kumar Guptaand Sudhakar Sengan (2022). *Journal of Database Management* (pp. 1-21).

[www.irma-international.org/article/measuring-the-determining-factors-of-financial-development-of-commercial-banks-in-selected-saarc-countries/311092](http://www.irma-international.org/article/measuring-the-determining-factors-of-financial-development-of-commercial-banks-in-selected-saarc-countries/311092)

### Blockchain and Financial Market Innovation

Vandana Mehrotraand Meena Bhatia (2022). *Applications, Challenges, and Opportunities of Blockchain Technology in Banking and Insurance* (pp. 128-150).

[www.irma-international.org/chapter/blockchain-and-financial-market-innovation/306458](http://www.irma-international.org/chapter/blockchain-and-financial-market-innovation/306458)

### Managing Temporal Data

Abdullah Uz Tansel (2009). *Handbook of Research on Innovations in Database Technologies and Applications: Current and Future Trends* (pp. 28-36).

[www.irma-international.org/chapter/managing-temporal-data/20685](http://www.irma-international.org/chapter/managing-temporal-data/20685)

### The Quality of Online Privacy Policies: A Resource-Dependency Perspective

Veda C. Storey, Gerald C. Kaneand Kathy Stewart Schwaig (2009). *Journal of Database Management* (pp. 19-37).

[www.irma-international.org/article/quality-online-privacy-policies/3402](http://www.irma-international.org/article/quality-online-privacy-policies/3402)

### A Roadmap on Updates

Fernando Zacarías Flores, Dionicio Zacarías Flores, Rosalba Cuapa Cantoand Luis Miguel Guzmán Muñoz (2009). *Database Technologies: Concepts, Methodologies, Tools, and Applications* (pp. 2261-2267).

[www.irma-international.org/chapter/roadmap-updates/8034](http://www.irma-international.org/chapter/roadmap-updates/8034)