



Hartmut Wedekind

Preface

Data management systems play the most crucial role in building large application systems. Since modern applications are no longer single monolithic software blocks but highly flexible and configurable collections of cooperative services, the data management layer also has to adapt to these new requirements. Therefore, within recent years, data management systems have faced a tremendous shift from the central management of individual records in a transactional way to a platform for data integration, federation, search services, and data analysis. This book addresses these new issues in the area of data management from multiple perspectives, in the form of individual contributions, and it outlines future challenges in the context of data management. These contributions are dedicated to *Prof. em. Dr. Dr.-Ing. E. h. Hartmut Wedekind* on the occasion of his 70th birthday, and were (co-)authored by some of his academic descendants.

Prof. Wedekind is one of the most prominent figures of the database management community in Germany, and he enjoys an excellent international reputation as well. Over the last 35 years he greatly contributed to making relational database technology a success. As far back as the early 1970s, he covered—as the first author in Germany—the state of the art concerning the relational model and related issues in two widely used textbooks “Datenbanksysteme I” and “Datenbanksysteme II”. Without him, the idea of modeling complex-structured real-world scenarios in a relational way would be far less developed by now. Among Prof. Wedekind’s academic achievements is the development of an alternative way for schema design based on the adaptation of the structure of natural languages in combination with the concept of abstraction. With the relational background in mind, he may be called the father of the language-based schema design approach, which is still feasible in tackling complicated schema design issues. He has covered this broad body of knowledge in eight textbooks so far, many of which have appeared in several editions and reprints.

The impact of Prof. Wedekind’s work, however, reaches beyond pushing the relational idea (better: fighting an almost religious war at a time when data management systems were dogmatically based on hierarchical or network data models). As a founding member of the *Gesellschaft für Informatik e.V.* (German Association for Computer Science) in 1969, he greatly contributed to the overall development of computer science as a science in Germany. Even more impressive, and not directly visible through publications, official posts, etc., is the person of Prof. Wedekind. His sharp and brilliant mind in combination with a generally open-minded character is simply fascinating. As a special guest at a database conference in 1998 in Capri, he was introduced with the words “Mr. Wedekind is always concerned about the WHY behind the HOW”—a very appropriate description.

Following one of his citations, “Etwas verstehen heißt verstehen, wie es geworden ist” (Prof. Schnädelbach: “Something to understand means to understand how it has evolved”), it is advisable to say a few words about Hartmut Wedekind’s vita.

Prof. Wedekind was born in 1935 in Bochum, Germany. In 1955, he began his studies in industrial engineering (‘Wirtschaftsingenieurwesen’) with a strong emphasis on mechanical engineering (‘Fachrichtung Maschinenbau’) at the Darmstadt University of Technology. After finishing his diploma in Darmstadt, he ventured out (1961/62) and earned his Master of Science in Engineering, emphasis on Operations Research, at the University of Berkeley, advised by George B. Dantzig. Back in Darmstadt, he finished his PhD thesis (Dr. rer. pol.) in 1963. Thereafter, Dr. Wedekind joined IBM in Sindelfingen and Frankfurt, where he was involved in application design and programming based on the IBM 1401/1440. Out of IBM, he earned the *Venia Legendi* through his habilitation, advised by Prof. Bussmann and Prof. Heinhold, 1967. One year later, he accepted a position as professor (for Business Studies) at the TU Darmstadt, where he initiated the now renowned Computer Science Department and became its first dean. During a stay at the IBM Research Lab in San Jose as a visiting scientist in 1972, he worked with Edgar F. Codd. The first ideas of the relational data model were discussed at that time. As he himself points out, it dawned on him that the relational idea was of fundamental importance and implied a revolution in storing and querying data sets—this was a terrific time! ... years before the notion of transaction was born.

In 1979, Prof. Wedekind changed his affiliation and moved to the University of Erlangen-Nuremberg, where he established a new data management group. His successful work is documented in numerous publications. Moreover, he managed a special research program (Sonderforschungsbereich) funded by the German Research Council for 12 years from 1987–1998. He officially stepped down from leading the database group in Erlangen in 2002. In recent years, he has spent even more time in elaborating the foundations of computer science from a philosophy-of-science point of view, and has presented important topics like abstraction, objects and metalanguage, or schemas and instances in academic lectures. He has advocated that these topics are essential for the field, although they are in general not treated in classical computer science courses and in lectures in other sciences that depend upon computer science. Furthermore, he is devoted to computer science education in schools, laying the foundations for the next-generation computer scientists. Nevertheless, he is currently still teaching at three different universities (Friedrich Alexander University of Erlangen-Nuremberg, Darmstadt University of Technology, and Friedrich Schiller University of Jena) and publishing papers in national as well as international journals and periodicals.

In more than 30 academic years, he has “produced” a large number of PhDs (see the pedigree of academic descendants below and in the appendix). As can be seen, his family is already in its fourth generation and the spirit of Prof. Wedekind—take nothing for granted, think around the corner, and always look behind the scenes—is propagated to all his academic descendants.

To honor Prof. Wedekind’s work in general and his 70th birthday specifically, we organized an international symposium held from June 30th to July 2nd, 2005 at the International Conference and Research Center for Computer Science, Schloss Dagstuhl. All contributions covered by the book were presented by their authors.

Structure of the Book

This book is logically divided into four parts, which are introduced briefly along with their associated articles.

Part I: Motivation and Modeling Issues

The first part of the book discusses the general perspective of the central role of databases in a modern information society. The contribution

Databases: The Integrative Force in Cyberspace,

authored by *Andreas Reuter*, demonstrates the shift of database technology from a central repository to a platform for transparent integration and efficient analysis of different heterogeneous data sources. Different classes of data are introduced, technological trends are explained, and a shape of future database systems is outlined. The general introduction is followed by an article written by *Bernhard Mitschang*—one of the first grandchildren of Prof. Wedekind—and his group. This discussion,

Federating Location-Based Data Services,

outlines the impact of location awareness on data management services in ubiquitous computing environments. Using the Nexus platform as an example, they point out the different requirements and characteristics for creating a federated data service. Different architectural approaches like LAV and GAV are analyzed in this context. The third piece in this part,

An Agent-Based Approach to Correctness in Databases,

was contributed by *Herbert Stoyan* and his group and discusses the approach of using agent technology to detect semantic inconsistencies, as an illustrative example referring to a database with genealogical data on historic persons belonging to European nobility.

Part II: Infrastructural Services

The second part of this Festschrift volume covers contributions with a focus on infrastructural services. The introductory essay reviews the history of server technology over the last 30 years. The article

Thirty Years of Server Technology—From Transaction Processing to Web Services,

penned by *Klaus Meyer-Wegener*, outlines how (application) server technology developed in recent decades and argues about current trends and future requirements. This overview is followed by a discussion of the various forms of caching in the Internet. Caching—in the sense of weak database replication—is a necessity for distributed applications within a totally connected world.

Caching over the Entire User-to-Data Path in the Internet,

authored by *Theo Härder*, underlines the need for caching Web objects in different proxy caches in the client-to-server path and explains how database caching based on cache constraints enables declarative query processing close to application servers, thereby speeding up the server-to-database path. This contribution is followed by two reports addressing the concept of messaging and notification services. The first of these two,

*Reweaving the Tapestry: Integrating Database and Messaging Systems
in the Wake of New Middleware Technologies,*

contributed by the working group of *Berthold Reinwald* at the IBM Almaden Research Center in San Jose, outlines the issue of adding messaging services to the query-based interaction model of classical database systems. This evolutionary approach is complemented by an almost revolutionary approach of building notification systems by exploiting existing database technology, as described in

Data Management Support for Notification Services,

by *Wolfgang Lehner*. The infrastructural part closes with a discussion on extending search capabilities in database systems to better support applications like full text search, Web page indexing, etc. The final piece,

Search Support in Data Management Systems,

by *Andreas Henrich*, outlines the requirements and presents the current state of the art concerning search methods.

Part III: Application Design

The third part of this book focuses on multiple facets of designing large applications. Application design from the proposed perspective falls into the category of “Programming in the Large,” which was always a main focus of Hartmut Wedekind’s academic work. The first article, entitled

Toward Automated Large-Scale Information Integration and Discovery,

and authored once again by the group of *Berthold Reinwald*, considers the problem of information integration as a single step towards a more global enterprise application integration. A more conceptual discussion on how to build large applications for a specific class of information systems is described in

Component-Based Application Architecture for Enterprise Information Systems,

by *Erich Ortner*. In a first step, this analysis explores the characteristics of the general component concept, followed in a second step by a discussion of tools and appropriate design methodologies to build enterprise-scale applications. The third contribution reviews the notion of workflows from a general perspective. In

Processes, Workflows, Web Service Flows: A Reconstruction,

Stefan Jablonski discusses the history and future developments of the area of workflow management systems and introduces a general model to specify business processes. The same area is addressed in the discussion of distributed execution of workflows in

Pros and Cons of Distributed Workflow Execution Algorithms,

penned by *Hans Schuster*. His paper states the problem and compares state-of-the-art techniques in this context. He also pinpoints major drawbacks and proposes adequate solutions. The section on application design is rounded out by

Business-to-Business Integration Technology,

written by *Christoph Bussler*, where he sketches the major issues in this area and discusses database extensions in order to fulfil the requirement of being an integrative platform even for cross-organizational, i.e., business-to-business processes.

Part IV: Application Scenarios

The final part of this book is dedicated to application scenarios with a strong emphasis on data management issues. The first two contributions exhibit a financial background. In

Information Dissemination in Modern Banking Applications,

authored by Peter Peinl and Uta Störl, an FX Trading application is used as an example to demonstrate real requirements with respect to state-of-the-art data management platforms. While this article focuses on the structural perspective, the following contribution,

An Intermediate Information System Forms Mutual Trust,

written by Dieter Steinbauer, outlines the process including the underlying information system necessary to decide on granting a credit to a customer within the SCHUFA environment. The next article, entitled

Data Refinement in a Market Research Applications' Data Production Process,

illustrates the data production process within the data-warehouse environment of GfK Marketing Services. The most distinctive feature of the application is that data are treated like industrial goods, i.e., highly customized data products are designed for individual customers; Thomas Ruf and Thomas Kirsche have written this illustrative report. The third application context covered within this series of application-oriented articles is devoted to health care. In

Information Management in Distributed Healthcare Networks,

Richard Lenz discusses the specific requirements of information management in such contexts and outlines major challenges and some solutions. The last but definitely not least contribution comes from Hans-Peter Steiert and has the title

Data Management for Engineering Applications.

In his analysis, the integration of data and processes in the engineering application area is discussed at different levels. Furthermore, soft factors of the integration problem, such as dependability management and data flow integration using patterns, are included in his considerations.

To summarize, this book provides a comprehensive overview of the state of the art and future trends in technology and applications centered around data management issues.

Kaiserslautern/Dresden
May 2005

Theo Härder and Wolfgang Lehner

Pedigree of Academic Descendants



