The Database and Information System Research Group at the University of Ulm

Peter Dadam, Wolfgang Klas

University of Ulm
Faculty of Computer Science
Dept. Databases and Information Systems
D-89069 Ulm, Germany
E-mail: {dadam, klas}@informatik.uni-ulm.de
WWW: http://www.informatik.uni-ulm.de/dbis

1. General Overview

The University of Ulm was founded in 1967 with focus on medicine and natural sciences. In 1989 the University established two new faculties: Engineering Sciences and Computer Science. This enlargement took place within the framework of the so-called *Science City Ulm*. In a joint effort, the State of Baden-Württemberg, industrial companies, the University, and the City of Ulm successfully established a research and development infrastructure at or nearby the university campus consisting of the university's research labs, university-related research institutes like the Research Institute for Applied Knowledge Processing (FAW), and industrial research and development labs, especially a large research center of Daimler-Benz AG.

Today, the Faculty of Computer Science consists of seven divisions (called 'departments'), each of which equipped with two professor positions:

- Theoretical Computer Science
- Artificial Intelligence
- Distributed Systems
- Databases and Information Systems
- Software Technology and Compiler Construction
- Computer Structures
- Neural Information Processing.

The Dept. of Databases and Information Systems (DBIS) became operational at the beginning of 1990 when Peter Dadam joined the faculty. He came from the IBM Heidelberg Science Center (HDSC) where he managed the research department for *Advanced Information Management* (AIM). At the HDSC he was working on advanced database technology and applications and contributed to the development of the AIM-P system (see [1]). The second professor position was first occupied by Marc Scholl, who belonged to the DBIS department from 1992 to 1994. In 1996 Wolfgang Klas joined the DBIS department as second professor. He came from the GMD Institute for Integrated Publication and Information Systems (IPSI) where he managed the research division *Distributed Multimedia Information Systems* and was working on advanced object-oriented database systems technology, interoperable database systems, and multimedia information systems.

At present, the DBIS team consists of the teaching and research assistants Thomas Bauer, Susanne Boll, Christian Heinlein, Clemens Hensinger, Erich Müller, Manfred Reichert, Birgit Schultheiß, the system engineer Rudi Seifert, the secretary Christiane Köppl, and the doctoral students Thomas Beuter and Anita Krämer.

In the following, we concentrate on the research and development work performed previously and presently in the research groups of Peter Dadam and of Wolfgang Klas. For references to Marc Scholl's work please visit http://www.informatik.uni-konstanz.de/dbis.

2. Database Technology and Advanced Database Applications

2.1 Database Technology

Most of the research in this area was performed from 1990 to 1995. To a certain degree it was a logical continuation of the research activities performed in the AIM-P project at the HDSC. The two major areas were

- *Query optimization and support of flexible storage structures in extended NF^2 DBMS*

  Today's object oriented DBMS as well as most of the research prototypes based on the NF^2, extended NF^2, or similar data models usually provide a (more or less) hard-wired mapping from logical to physical storage structures. The research performed in this field concentrated on the issues of (a) achieving a higher degree of data-independence by supporting a variety of different physical storage structures for a given logical object structure, (b) finding rules for the utilization of (path) indexes during query processing, and (c) developing appropriate query optimization and execution techniques to utilize these features.

- *Support of variant structures in relational and NF^2 relational database management systems*

Today’s relational, object-oriented DBMS, and research prototypes for complex object support do not provide appropriate mechanisms to adequately support variant data structures. Neither null-values nor attribute inheritance in object-oriented data models do really solve this problem. The challenge was to develop a generalized, extended relational data model, to redefine the relational algebra operators accordingly, and to understand the consequences for query execution (cf. [3] for more details).

2.2 Advanced Hospital Information Systems

We consider advanced hospital information systems as being one of the most challenging and interesting application areas for databases, workflow technology, and software development in general. Very often, especially in the context of university hospitals, we find rather decentralized organizational structures with (at least partially) decentralized software development and maintenance leading to "information islands". We find personnel working under (partially extremely) high time pressure, who must often make important decisions about patient treatment within a short period of time. We find strictly predefined medical procedures which have to be obeyed on the one side, and on the other side find many cases and circumstances in which ad-hoc deviations from these procedures are mandatory (by nature this probably happens much more frequently than in any other application area). – The DBIS department, therefore, has been engaged in clinical projects almost from the very beginning.

From 1992 to 1995 the DBIS department was leader of the large interdisciplinary research project *Open Clinical Database and Information System for the Integration of Autonomous Subsystems* funded by the State of Baden-Württemberg. Within this project several clinics (internal medicine, clinical chemistry, anesthesiology), the hospital computing center, and the DBIS department worked closely together to identify the technological requirements with respect to adequate clinical information systems. The (initial) goal of this project was the development of concepts for the provision of information and application services in a heterogeneous, open systems environment (see [4]) with existing information islands, decentralized software development and thus decentralized provision of electronic services.

During the project it turned out, however, that although improved access to patient data is very important, it will not sufficiently solve the core problems in the clinical routine. The reason is that today physicians are already confronted with a massive load of data which have to be intellectually processed and structured. As this usually has to be done within a rather short period of time this leads to many kinds of mistakes. An advanced computer-based system should therefore not only provide on-line access to patient data, but should also relieve medical personnel from tracking the state of medical activities like processing of a medical order, and from caring about the timely collection of their results.

In the sequel, we have therefore concentrated on concepts for process-oriented (clinical) application systems which actively support the medical personnel in keeping track of events, in scheduling and in coordinating tasks, and in cooperating with other clinical staff. The challenge was on the one side to understand in-depth all characteristic types of processing, the organizational structures, the kinds of exceptions which may occur in clinical processes, and the adequate reaction on such events, to learn how computer support can be smoothly integrated in the daily routine work, and how adequate man-machine interfaces for clinical environments have to look like. – The challenge on the other side was to find the adequate technological basis for supporting these requirements (see Sect. 3).

On the application side we are working into two directions: "Smooth assistance" of routine work by knowledge-based components on the one side, and business process reengineering and workflow management (the larger activity) on the other side. With "smooth assistance" we mean that in clinical application environments the computer may only make suggestions but is never allowed to make final decisions. This is especially true when decisions about patient treatment have to be made (because of moral but also of legal reasons). In order to understand how such an assistance could look like, we have developed a component to assist the physician in selecting tests and medical interventions based on medical guidelines which suggests (further) tests and treatments for a given indication and on outcomes obtained so far (see [2]). – Similar rules could be used in principle to describe inter-workflow dependencies on a logical level. These rules could be used as a basis for generating the input for an inter-workflow execution monitor (see Sect. 3.3).

At the beginning of 1996, we started the two-years project *Using Workflow Management Systems for Clinical Applications* in cooperation with the Women's Hospital of the University of Ulm and the workflow division (WorkParty development) of Siemens-Nixdorf AG. The goal is to elaborate a complete "process map" for this hospital, and to evaluate to which extent today's workflow technology – especially the workflow management system (WfMS) WorkParty – is (already) able to meet the requirements of clinical workflow applications.

Within this project we have thoroughly analyzed the organizational structures and deeply analyzed all relevant (business) processes of this hospital. Some of them have been redesigned and optimized, which has partially led to significant improvements. To gain concrete implementation and usability experience we have implemented the complete business process (including the integration of foreign systems) for the division "day clinic" (which performs minimal invasive surgery) of this hospital based on WorkParty. We have completely replaced the standard user interfaces of the underlying WfMS by own application-oriented end-user interfaces, and we have also extended the capabilities of the WfMS by the provision of notification services (making the server an "active" component), the support of time, and "semantic rollback" across several steps (see also Sect. 3).

We expect that the clinical domain will continue to be one of our major areas of application-oriented research and a good test bed for our research work in cooperative information systems and workflow technology.

2.3 Concurrent Engineering

Since 1993 we have closely cooperated with the Daimler-Benz Research Center in Ulm in the field of Concurrent Engineering. The ultimate goal is to improve the product development process, i.e., to shorten development times, to maintain or improve product quality, and to reduce costs. Within this process we are contributing to the identification and, where necessary and appropriate, to the development of the adequate technological basis to perform this task.

3. ADEPT♣: Cooperative Information Systems and Workflow-Management

3.1 Dynamic Changes

One of the most severe weaknesses of today's workflow technology is the absence or very limited ability to deviate (in a controlled and secure way) from the pre-planned process template at run-time. For many clinical processes (except some simple ones) it is usually either not possible or at least not cost-effective to put all possible deviations from the "standard plan" as possible execution sequences into the process template. We are therefore working towards workflow technology which allows to deviate from the pre-modeled process template (skipping of steps, going back to previous steps, inserting new steps, etc.) in a secure and safe way. That is, the system guarantees that all consistency constraints (e.g., no cycles, no missing input data when a task program will be invoked) which have been ensured prior to the dynamic ("ad hoc") modification of the process instance are also ensured after the modification (see [5], [6]).

3.2 Temporal Aspects

Today's WfMS usually support the concept of time rather rudimentarily. Deadlines for the execution of steps can be specified, but the temporal dependencies between them are usually not known to the system. Thus, the consequences for missing a certain deadline with respect to subsequent steps are not directly inferable. The experiences with clinical applications (see Sect. 2.2) have led to the conclusion that the support of time and time-dependencies are very important features of a WfMS to become applicable for a broader range of applications. By doing so, WfMS are going to provide functionality usually found in calendar and scheduling systems. – This functionality especially becomes important in the context of supporting dynamic changes as described in the previous section.

3.3 Inter-Workflow Dependencies

Many real-world tasks – not only in the clinical context – are interrelated in the one way or another. The tasks "endoscope examination of the stomach" and "ultrasound examination of the stomach", for example, are interrelated in at least two ways: Normally, they are performed in different examination rooms and can therefore not be executed for the same patient at the same time. In addition, they also interfere with one another. If the endoscope examination (which requires to fill the stomach with air) is performed at first, the ultrasound examination cannot be performed on the same day. At first glance, it would therefore make sense to combine such interrelated tasks at the modeling level into one single workflow template (process template) which would, in principle, open the possibility to obey all the dependencies by ordering the steps within the overall workflow accordingly. However, this approach causes problems for several reasons: At first, the two tasks need not to be initiated at the same time. That is, each possible interleaving of steps may occur at execution time and would have to be taken into account at the modeling level. In addition, it is not sufficient to consider the tasks in pairs. Every combination of interfering tasks – with all possible interleaving – would have to be considered. Obviously, this would lead to a huge number of workflow templates, each of which very complex and hard to maintain. Therefore, mechanisms have to be developed which allow to express inter-workflow dependencies and to monitor and steer the execution of such workflow instances at run-time.

3.4 Further Activities

Further activities in this area are falling into the categories:

- *Component-based software development*
- *Environments for large-scale workflow executions*
- *Workflow scheme evolution*
- *Modeling concepts*
- *Man-Machine interfaces*
- *Architectural issues and ADEPT implementation*

♣ ADEPT stands for <u>A</u>pplication <u>D</u>evelopment Based on <u>E</u>ncapsulated <u>P</u>remodeled Process <u>T</u>emplates.

4. Multimedia Information Systems

4.1 Multimedia Repositories

Our work in the area of multimedia information systems focuses on the integrated management and presentation of multimedia data by a DBMS. The overall approach taken is driven by a project aiming at the development of a network-based and database-driven multimedia information system for doctors, medical lecturers, students, and patients. The application requirements are provided by the University Hospital of Ulm, its Cardiological Clinic, Surgical Clinic, and the associated Rehabilitation Hospital. In this context, we are developing concepts and prototypical implementations of a multimedia data repository that supports

- efficient management of static multimedia material for courses in the domain of heart surgery,
- dynamic multimedia presentations including photographs, slides, X-ray images and findings, video recordings of heart surgeries, and audio annotated material,
- interactive case-based training and teaching material,
- patient-oriented information.

The conceptual and technical challenges arise from the approach to seamlessly integrate flexible dynamic multimedia presentation services with the multimedia repository. We are concerned with the database-driven generation of multimedia presentations depending on output channels, e.g., university campus-based or home-based channels, on the actual performance of infrastructure as well as on the presentation quality and content requested by the consumer. This calls for a presentation-neutral representation of multimedia material in the database, e.g., representing text as SGML documents (e.g., [7]), providing images, audio, and video data in different presentation quality, and allowing to store modular presentation plans independently of the final presentation format. The system has to support the dynamic composition of fragments of such presentation plans according to user requests. That is, the repository will contain pre-orchestrated "presentation beans" which can be composed to final presentations on the fly. The quality of a presentation may still be a parameter for such a composition of presentation beans and will be determined by the output channel chosen by the user. For example, a high quality presentation dynamically composed during a seminar at the university campus should be available for students at home although they do not need the high quality of videos or images at home. This obviously requires a configurable presentation execution environment as the front end of the repository. The repository has to provide rich modeling capabilities including the representation of metadata for efficient indexing of multimedia material [9] and it has to support a variety of standards including, e.g., the medical standard DICOM (Digital Imaging and Communications in Medicine).

The approach taken in this project aims at the provision of a generic presentation service by the multimedia repository [8, 10, 11, 12, 13]. It is based on the AMOS project which initially started at GMD-IPSI, and it makes use of object-relational database technology. The prototype is being implemented in Java.

4.2 Internet-based Information Systems

In the context of an international project sponsored by the European Commission we are currently prototyping a database-driven electronic commerce system, the Electronic Market Place (EMP). With this system producer and consumer in rural and remote regions can participate in the market at the same conditions as participants in the market in urban regions. The electronic market place allows to execute business transactions, i.e., to offer and to buy goods and services without any intermediary and wholesalers. The market trades not only arbitrary goods but offers as well services like financing and logistics. These services are integrated in the traditional buying and selling services. All product and service data is safely and consistently managed by a central instance, the EMP server. The traders connect to the EMP server and place offers or retrieve offers from there.

The realization of the project is based on existing infrastructure such as WWW and Internet. But it will still be possible to participate with traditional communication channels such as fax, phone, and mail. The major issues addressed in our group are the design of the database server which handles all kinds of media objects, the design and integration of the EMP services with the database server, and the realization of these components in Java.

5. Other IS-related Research

The group of Michael Weber, Dept. for Distributed Systems (see http://www-vs.informatik.uni-ulm.de), pursues research closely related to the DBIS activities. They focus on system support for cooperative applications and collaborative scenarios. Specific research projects are:

- *Middleware for integrated cooperative systems*
  The concepts of brokerage and trading in distributed systems have been extended to seamlessly bind asynchronous (e.g., workflow management) and synchronous (e.g., desktop conferencing) CSCW systems. This approach combines techniques known from distributed object computing with methods from agent technologies. These are augmented by concepts tailored for multi-user, cooperative applications. Proofs of these concepts have been demonstrated in real life scenarios taken from civil authorities. A dedicated trial is currently being conducted with building constructors and architects.

- *Collaboration support for teleteaching*

In a joint project with members of various faculties a teleteaching framework is being developed. While other partners participate in the role of content providers, the group of Michael Weber is enhancing the framework by customizable and adaptive collaboration support. A range of components including audio/video conferencing, discourse system, help desk and lightweight workflow management is offered. All components can be linked to the content using Java applet technology.

- *WWW-based call center for sensitive medical applications*
  This generic call center provides Internet remote access to patient databases being set up for specific diseases or syndromes. A help desk extends the plain database retrieval by means of tele-consultation. Major research issues are security and privacy in such Internet-based applications.

- *Distributed multimedia information systems*
  Together with media authors and screen designers a municipal information system has been developed. This system uses an ATM broadband network in town and slow lines to connect to the Internet. Software architectural issues concerning the diverse network capabilities while ensuring fast response times and artistic screen design have been the focus of this work.

6. Infrastructure

For our teaching as well as for our research we consider it very important to teach not just concepts and to work not only at a conceptual level, but also to have hands-on experience with existing systems. We, therefore, operate a large variety of systems in our department. At present we have installed the following commercial systems:

Hardware: Sun, RS/6000, PCs

Operating Systems: Solaris, AIX, Novell, OS/2, Win 3.11, Windows NT

Database Management Systems: Oracle, DB2, O_2, ObjectStore, Illustra/Informix with various multimedia and web extensions

Modeling Tools: Aris-Toolset, Aeneis, Bonapart, LeuSmart, ProMod, Statemate

Workflow Management and Groupware Systems: Aris-Workflow, FlowMark, Groupflow, Lotus Notes, ProMInanD, WorkParty

Other Software: e.g., CICS/6000, openUTM, Encina, DCE

7. References

[1] Dadam, P., Linnemann, V.: Advanced Information Management (AIM): Advanced Database Technology for Integrated Applications, IBM Systems Journal, Vol. 28, No. 4, 1989, pp. 661-682.

[2] Heinlein, C.; Kuhn, K.; Dadam, P.: Representation of Medical Guidelines Using a Classification-Based System. Proc. CIKM '94, Gaithersburg, Maryland, Nov./Dec. 1994, pp. 415-422.

[3] Kalus, C., Dadam, P.: Flexible Relations - Operational Support of Variant Relational Structures, Proc. VLDB '95, Zürich, Sept. 1995, pp. 539-550.

[4] Kuhn, K.; Reichert, M.; Nathe, M.; Beuter, T.; Dadam, P.: An Infrastructure for Cooperation and Communication in an Advanced Clinical Information Systems. Proc. 18th Ann. Symp. on Computer Applic. in Medical Care 1994, Washington, 1994.

[5] Reichert, M.; Dadam, P. ADEPT$_{flex}$ – Supporting Dynamic Changes of Workflows Without Loosing Control. Univ. of Ulm, Faculty of CS, Res. Rep. No. 97-07, April 1997.

[6] Reichert, M.; Dadam, P.: A Framework for Dynamic Changes in Workflow-Management Systems. Proc. Int'l Workshop on Database and Expert Systems Applic., Toulouse, France, Sept. 1997, pp. 42-48.

[7] Böhm, K., Aberer, K., Klas, W.: Building a Hybrid Database Application for Structured Documents, Journal on Multimedia Tools and Applications, 5, 1997, Kluwer Academic Publishers, pp. 275-300.

[8] Boll, S., Klas, W., Löhr, M.: Integrated Database Services for Multimedia Presentations. In: S.M. Chung (ed.), *Multimedia Information Storage and Management.* Kluwer Academic Publishers, 1996.

[9] Boll, S., Klas, W., Sheth, A.: Overview on Using Metadata to Manage Multimedia Data. In W. Klas, A. Sheth (Eds.): Managing Multimedia Data: Using Metadata to Integrate and Apply Digital Data, McGraw-Hill, 1997.

[10] Boll, S., Wäsch, J.: A Java Application Programming Interface to a Multimedia Enhanced Object-Oriented DBMS. Proc. First Int'l Workshop on Persistence and Java (PJ1), Glasgow, Scotland, Sept. 1996.

[11] Klas, W., Aberer, K.: Multimedia and its Impact on Database System Architectures. In: P. Apers, H. Blanken, M. Houtsma (eds.): *Multimedia Databases in Perspective,* Springer Verlag Berlin/Heidelberg, 1997.

[12] Thimm, H., Klas, W., Cowan, C., Walpole, J., Pu, C.: Optimization of Adaptive Data-Flows for Competing Multimedia Presentational Database Sessions, IEEE Int'l Conf. on Multimedia Computing and Systems, June 1997, Ottawa, Canada.

[13] Vazirgiannis, M., Boll, S.: Events in Interactive Multimedia Applications: Modeling and Implementation Design, IEEE Int'l Conf. on Multimedia Computing and Systems, June 1997, Ottawa, Canada.