

Commenced Publication in 1973

Founding and Former Series Editors:

Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

Editorial Board

David Hutchison

Lancaster University, UK

Takeo Kanade

Carnegie Mellon University, Pittsburgh, PA, USA

Josef Kittler

University of Surrey, Guildford, UK

Jon M. Kleinberg

Cornell University, Ithaca, NY, USA

Friedemann Mattern

ETH Zurich, Switzerland

John C. Mitchell

Stanford University, CA, USA

Moni Naor

Weizmann Institute of Science, Rehovot, Israel

Oscar Nierstrasz

University of Bern, Switzerland

C. Pandu Rangan

Indian Institute of Technology, Madras, India

Bernhard Steffen

University of Dortmund, Germany

Madhu Sudan

Massachusetts Institute of Technology, MA, USA

Demetri Terzopoulos

New York University, NY, USA

Doug Tygar

University of California, Berkeley, CA, USA

Moshe Y. Vardi

Rice University, Houston, TX, USA

Gerhard Weikum

Max-Planck Institute of Computer Science, Saarbruecken, Germany

Matthias Hemmje Claudia Niederee
Thomas Risse (Eds.)

From Integrated Publication and Information Systems to Virtual Information and Knowledge Environments

Essays Dedicated to Erich J. Neuhold
on the Occasion of His 65th Birthday



Springer

Volume Editors

Matthias Hemmje
FernUniversität Hagen
Universitätsstr. 1, 58097 Hagen, Germany
E-mail: Matthias.Hemmje@fernuni-hagen.de

Claudia Nederee
Thomas Risse
Fraunhofer IPSI
Dolivostr. 15, 64293 Darmstadt
E-mail: {claudia.nederee,thomas.risse}@ipsi.fraunhofer.de

Library of Congress Control Number: 2004118429

CR Subject Classification (1998): H.3, H.4, H.2, H.5, I.2

ISSN 0302-9743

ISBN 3-540-24551-0 Springer Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.

Springer is a part of Springer Science+Business Media

springeronline.com

© Springer-Verlag Berlin Heidelberg 2005
Printed in Germany

Typesetting: Camera-ready by author, data conversion by Markus Richter, Heidelberg
Printed on acid-free paper SPIN: 11387336 06/3142 5 4 3 2 1 0

Preface

Matthias Hemmje

FernUniversität in Hagen, Department of Computer Science,
Universitätsstraße 1, D-58097 Hagen, Germany
Matthias.Hemmje@fernuni-hagen.de

This book is dedicated to Erich J. Neuhold, Director of Fraunhofer's Integrated Publication and Information Systems Institute (IPSI) and Professor of Computer Science at the Technical University of Darmstadt on the occasion of his 65th Birthday in January 2005.

Erich Neuhold's primary research and development interests are in heterogeneous multimedia database systems, in peer-to-peer- and Grid-based information environments, in Web-based distributed information technologies, in persistent information and knowledge repositories, and in content engineering. Within his interest in content engineering special emphasis is given to all technological aspects of the information and publishing process, i.e., the knowledge value chain that arises for the creation of digital products and services in Web application contexts. Besides his main areas of scientific work, Erich Neuhold also guides research and development in user interfaces including information visualization, computer-supported cooperative work, ambient intelligence, mobile and wireless technology, and security in Web application areas like e-learning, e-commerce, e-culture and e-government.

Building on its long-standing scientific expertise and reputation Erich Neuhold's Fraunhofer IPSI aimed at supporting information-, content- and knowledge-intensive applications by focusing on research and development towards the achievement of the following vision: the efficient and flexible implementation of technologies supporting the creation and effective utilization of virtual information and knowledge environments where such environments are Web-based distributed information systems that enable the corporate as well as the personal and cooperative acquisition, management, access, distribution and usage of information, content and explicitly encoded knowledge objects.

In accordance with the unifying vision and continuing research interest of Erich Neuhold and in correspondence with his scientific work, this book presents a selected number of invited papers by leading international researchers in the area, addressing a broad variety of related topics. It therefore reflects the breadth and the depth of Erich Neuhold's scientific work advocating the synergistically integrated approach towards supporting digital information and knowledge production on the one hand together with its corresponding access and usage support on the other hand.

As the introductions to the first, second, and fourth parts of this book are provided in the later opening sections by Tamer Özsu and Keith van Rijsbergen, I would now like to introduce the content of the other parts.

The third part is dedicated to issues related to Securing Dynamic Media Content Integration and Communication. In the first paper of this part, Ralf Steinmetz (who

successfully headed the Fraunhofer IPSI together with Erich Neuhold for several years and is currently still his colleague as the Professor for Multimedia Communications at the Technical University of Darmstadt) and Oliver Heckmann, address the subject of video distribution with peer-to-peer file-sharing networks. They introduce one of the most successful peer-to-peer file-sharing applications – it was especially successful in Germany – and discuss traffic characteristics and user behavior in distributing video files based on the results of their experiments and measurements. In the next paper, Jana Dittmann, former division head at Erich Neuhold’s Fraunhofer IPSI and now Professor for Multimedia and Security at the Otto-von-Guericke-University in Magdeburg, and Martin Steinebach, currently division head at IPSI, describe the design of a framework for media production environments, where mechanisms like encryption, digital signatures and digital watermarking help to enable a flexible yet secure handling and processing of the content. Finally, Andreas Meissner, also division head at IPSI, and Wolfgang Schönfeld propose a heterogeneous and flexible communication platform that complies with the reliability and coverage requirements of the Public Safety Organization.

The fifth part is titled Visualization – Key to External Cognition in Virtual Information Environments. In the first paper of this part Tiziana Catarci and her colleagues Enrico Bertini, Lucia Di Bello and Stephen Kimani of the University of Rome “La Sapienza” discuss how visualization has been and can be used to support interaction with the digital library environment. In the next paper, Gerald Jäschke, Piklu Gupta and I introduce an approach to design a declarative language for defining information visualization applications (IVML). Next to this, Ben Shneiderman and Jinwook Seo of the University of Maryland present a generalized knowledge integration framework to better understand the importance of an item or cluster by using multiple interactively coordinated visual displays and show examples of hierarchical clustering of gene expression data, coordinated with a parallel-coordinates view and with the gene annotation and gene ontology. The part is completed by Maria Francesca Costabile and Paolo Buono of the University of Bari presenting a visual strategy that exploits a graph-based technique and parallel coordinates to visualize the results of algorithms for mining association rules which helps data miners to get an overview of the rule set they are interacting with and enables them to deeper investigate inside a specific set of rules.

The sixth part takes us From Human Computer Interaction to Human Artefact Interaction. In the first paper of this part, Norbert Streitz, currently division head at Erich Neuhold’s Fraunhofer IPSI, outlines his approach to designing environments that exploit the advantages of real objects and at the same time use the potential of computer-based support by moving from human-*computer* interaction to human-*artefact* interaction. In the next paper, Peter Tandler and Laura Dietz (also members of Fraunhofer IPSI) explain how to support ubiquitous computing by enabling an extended view on sharing which seamlessly integrates the view of “traditional” CSCW and additionally incorporates ubiquitous, heterogeneous and mobile devices used in a common context. In the following paper, Jörg Haake, formerly PhD student of Erich Neuhold and now a Professor of Computer Science at the FernUniversität in Hagen, and his wife Anja Haake (also a former PhD student of Erich Neuhold) together with Till Schümmer introduce a metaphor and user interface for managing

access permissions in shared workspace systems. The part is then completed by the contribution of Jose Encarnação (a colleague of Erich Neuhold as a Professor of Computer Science at the Technical University of Darmstadt and as the Director of the Fraunhofer Institute for Computer Graphics (IGD) in Darmstadt) and his research collaborator Thomas Kirste; they introduce the new ICT paradigm of ambient intelligence, its concepts, system architectures, and enabling technologies.

Finally, the seventh part introduces Application Domains for Virtual Information and Knowledge Environments. In the first paper of this part, Thomas Kamps, a former PhD student of Erich Neuhold and currently a division head at the Fraunhofer IPSI, together with Richard Stenzel, Libo Chen and Lothar Rostek discusses an approach that shows how a mix of information extraction and classification methods can be used to automatically set up and update a network of business objects serving as a corporate memory index. In the next contribution, Hans-Jörg Bullinger, President of the Fraunhofer-Gesellschaft, Corporate Management and Research, and his research colleagues Alexander Slama, Oliver Schumacher, Joachim Warschat and Peter Ohlhausen of the Fraunhofer Institute for Industrial Engineering (IAO) in Stuttgart explain how ontologies can be applied to providing a model for the computer-supported representation of innovation projects and to enabling computer-aided identification of time-consuming constellations, and can therefore serve to provide a basis for achieving innovation-acceleration knowledge to ensure faster and better innovations. In the following paper, Claudia Niederée, Avaré Stewart and Claudio Muscoguri (all of the Fraunhofer IPSI) and I discuss approaches for both creating context awareness as well as supporting specific working contexts by means of virtual information and knowledge environments. In the next paper, Patrick Baudisch, a former PhD student of Erich Neuhold and now a researcher at Microsoft Research, together with Lars Bruecker presents a recommendation system that provides users with personalized TV schedules. The final contribution of this part is by Junzhong Gu and presents an approach to integrating consumer electronics (e.g., TV sets, DVD players, HiFi sets, etc.), mobile multimedia devices, as well as traditional computing devices (PCs, laptops, printers, scanners, etc.) into a complete application solution.

I would like to continue by thanking a number of people who strongly supported the timely creation of this book.

First of all, I am grateful to all the authors who contributed efficiently in a very short period of time to provide their high-quality scientific content for this book in order to honor Erich Neuhold's 65th birthday.

Furthermore, I am grateful to Wolfgang Glatthaar, Chairman of Fraunhofer IPSI's Kuratorium (advisory board), for supporting this book with a statement honoring the occasion from Erich Neuhold's research institute's point of view.

Next, I am grateful to Tamer Özsu and Keith van Rijsbergen for supporting the book by providing introductions to selected parts of the book on the one hand, on the other hand for contributing laudations for Erich Neuhold on behalf of the scientific communities that Erich Neuhold has been operating in, collaborating with and serving during his outstanding academic career.

With respect to coping with the editorial work for this book, I am grateful to my co-editors Claudia Niederée and Thomas Risse who proved themselves as reliable collaborators in what turned out to be a more challenging and demanding endeavor

than we originally expected, especially in already tough times of organizational demand and change.

Concerning assistance in the preparation of this book, I am grateful to Anja Rieber, a secretary of excellence at Fraunhofer IPSI, who provided invaluable office administration back-end support as well as a competent and reliable communication and coordination front end for the whole preparation process.

Giving credit to the organizational background support, I am grateful to Dieter Böcker, Deputy Director of the Fraunhofer IPSI, and Emil Wetzel, Head of Administration of Fraunhofer IPSI, for providing me with the opportunity to take over responsibility for this book on the one hand, and on the other hand for providing the formal approval and support to use IPSI's infrastructure to prepare it.

Finally, I am grateful to Alfred Hofmann of Springer's Lecture Notes in Computer Science and his whole team for providing excellent publication support.

The last paragraphs of my preface shall be addressed and dedicated to Erich Neuhold himself.

First, I would like to state at this point that it has been a pleasure for me to work on this book honoring Erich's life-time achievements on the occasion of his 65th birthday as it has been taking me back to many memories about our collaboration history since 1991.

Furthermore, I would like to congratulate and thank him especially with respect to the many career opportunities that his research and management work as well as his personal guidance have opened up and continuously provided for many of the contributors to this book, including myself.

All of us who share such a "direct collaboration experience" with Erich know that working with him is a VERY demanding activity in many dimensions. However, it still is and has always been rewarding on many levels, too! Therefore, I would not like to have missed it. This holds especially true because I experienced Erich not only as a researcher who always strives for excellence but also many times and especially in tough situations as a man with a high degree of professionalism, integrity, and – most important – dedication, loyalty, and reliability with respect to his responsibility for taking care of the personal welfare of his staff members and the further development of their professional future.

I would like to conclude by wishing Erich Neuhold and his family all the best for the years to come. Personally, I am looking forward to many more years of interesting and fruitful collaboration.

December 2004

Matthias Hemmje

Organizational Laudatio and Personal Note

Wolfgang Glatthaar

Chairman of the Board of Trustees of Fraunhofer IPSI

In spite of the rapid growth of knowledge and information and a variety of usage scenarios on the World Wide Web, the goal of delivering the “right answer” to a question universally available remains a goal which is yet to be attained. According to Erich Neuhold’s research agenda, this “right answer” needs to be appropriate in its depth, context and timeliness, and should be available everywhere and adhere to social and legal conventions. This goal determines the basis for demanding and multi-faceted research topics. The application spectrum ranges from the repurposing of publishing products, and classroom-based and mobile distance learning, through to automatic content generation for mobile phones, handheld computers or navigation systems, and more. The Fraunhofer Integrated Publication and Information Systems Institute conducts research and applications development in these areas.

IPSI is a research institute that was founded in 1987 by the German National Research Center for Information Technology (GMD). Since its inception it has been under the direction of Prof. Dr. Erich J. Neuhold, who also leads the research and teaching group with the same focus in the Computer Science Department at the Technical University of Darmstadt. The close links with the university are reinforced through the teaching activities there of IPSI researchers. Over 100 people currently work in research at IPSI, of whom about 20% are from abroad.

Being responsible for continuing IPSI’s success in the future, Erich Neuhold and his research and management team started a transition phase in the first years of the new millennium. Most importantly, on July 11, 2001 IPSI was one of the eight research institutes of the German National Research Center for Information Technology (GMD) that became members of the Fraunhofer-Gesellschaft. Through this merger the Fraunhofer-Gesellschaft now encompasses 17 institutes out of 58 with about 3000 scientists doing applied research in the area of computer science and information technology. According to Erich Neuhold’s own statements, IPSI is still today in this transformation process, permanently assessing risks and challenges, and taking measures to adapt to the changed environment. Some of these consequences relate to the changes in the research paradigm – from basic research towards application-oriented and contract research – while others relate to rather mundane issues like building up the institute’s own marketing and sales processes.

From Erich Neuhold’s point of view, research at IPSI has never happened in an ivory tower. IPSI has always tried out new ideas and solutions on real-world problems. Its prime mechanism for reaching out to the market and turning research results into products has been the foundation of spin-offs. Due to Erich Neuhold’s vision and engagement, IPSI will continue to be an incubator of ideas in the years to come. Furthermore, many former employees of IPSI have started university careers in Germany

and abroad. Building on these contacts, IPSI has constantly used this opportunity to integrate its alumni in cooperation projects.

Under the leadership of Erich Neuhold, Fraunhofer IPSI in Darmstadt has been successfully developing user-centered systems that flexibly, effectively and efficiently respond to end-user needs in the application areas of electronic publishing and dissemination of information, collaborative work, and learning at common locations or over distances. The systems are formed using novel concepts of information enhancement and of information and knowledge management. Furthermore, IPSI investigates and develops high-quality software solutions for computer-supported collaborative work, electronic publishing and lifelong learning in real and virtual environments. IPSI's research areas to a high degree reflect Erich Neuhold's personal research interests, and therefore include digital libraries and information systems, publication tools with underlying database layers, distributed editing environments for collaborative maintenance of large collections of data, media and document management, and the more loosely related fields of knowledge management and cooperative learning systems, and security and services for mobile communication.

Values can only be preserved through change – so said Richard Löwenthal, the Anglo-German publicist and political scientist, in the twentieth century. Since its incorporation into the Fraunhofer Corporation in summer 2001, IPSI has been engaged in putting this insight into practice. During rigorous portfolio discussions and restructuring measures – in the course of which many favorite subjects and group approaches had to be sacrificed to a focus on areas of expertise more relevant to industry – this policy was amended. In this process, departments were analyzed down to group level, and in some cases were reconstituted – all with the thoroughgoing cooperation of many staff members, who, at the same time, had to carry on their usual work involving research, EU and BMBF (German Federal Ministry of Education and Research) projects, teaching, and industrial business. While this was proceeding, a new "Marketing & Sales" Department was formed. As a consequence, many new marketing instruments were introduced, followed by more rigid instruments for financial and sales control. However, all these achievements do not mean that the change process is now complete. As the Fraunhofer Institute operates in the area of research, it will never, and must never, become like a sort of screw factory, where everyone can sell everything – it is only natural that the concept of change has to be omnipresent at all times, by definition.

Besides managing IPSI during its GMD times as well as during the initial phases of its transformation into a Fraunhofer Institute, Erich Neuhold has always provided flexibility on the management level as well as on the administrative and operational staff level, supporting not only technologically but also aesthetically motivated and driven endeavors. Besides regularly hosting art exhibitions in the institute's building, for example during open house phases, he supported and initiated a set of artistically driven activities for the IPSI. Besides more or less hybrid IT R&D projects serving both worlds, like the Virtual Gallery projects resulting in cyber:gallery – a system for exhibiting and selling art on the Internet that enables the perception of art sales exhibitions in a VR-technology-driven Web-based application solution – there has also been room for supporting, for example, real and virtual theatre projects.

One such project originated at the beginning of 1999 when the Richard Wagner Forum in Graz offered the opportunity to submit projects for the staging of Richard Wagner's *Parsifal*. One option of the competition was to submit a cyberstaged version of *Parsifal* with an emphasis on the use of new technologies within the presentation. Supported by Erich Neuhold and IPSI, such a contribution from the IPSI staff member Christoph Rodatz was successful in the contest. During the project, traditional theatre architecture linked the audience with the stage through an audiovisual window. The stage was separated from the auditorium using a closed iron curtain. Thus the audience was able to move in a free way between the auditorium and the stage, and therefore the audience was in effect operating as a "hiker" between the two perception dimensions. Furthermore, the auditorium was transformed into an acoustic space of experience. Here *Parsifal* could be heard in a concerto form as a live transmission from the stage, except that, apart from the conductor, the performers, the orchestra and the choir were not present. Similarly, the stage area was transformed into a stage museum space, where the singers and musicians performed. The conductor's signal for the singers' entry was communicated via video. Each individual sound was sent via microphones to a sound panel. From there the entire "virtual soundscape" was transmitted to the auditorium. This approach meant that during the entire staging the orchestra and the choir could operate while still being in their rehearsal rooms, each equipped with microphones and at the same time quasi-"exhibited" in 10 vitrines that were located on the borders of the revolving stage. Via video screens and headphones the actors and singers were connected with the conductor and the entire soundscape experience system. Within so-called picture spaces there were 10 separate exhibition areas that could be entered via the revolving space. Furthermore, *Parsifal* was the topic of associated installations that enabled a more open access to Wagner's opera. With respect to the orchestra pit, the installation hosted video screens that displayed the movement of the live performance produced by each group of musical instruments transmitted from the orchestra and choir rehearsal rooms. The positioning of the video screens followed the specific pattern of the orchestra's natural position. Even though the chronology of the entire opera was not changed, an individual member of the audience could experience the entire staging in a fragmented and individualized manner if he or she decided to become a "hiker" during the performance. Using this approach, the audience could physically participate in the opera, and this had a tremendous impact on the dramatic degrees of freedom available within the director's interpretation of the *Parsifal* plot, bearing in mind that Wagner conceived of a "Gesamtkunstwerk". The cyberstaging of *Parsifal* demonstrated the influence of new technologies on society's perception, and the transformation that could be brought about in the theatre and in the arts in general.

The first time I met Erich Neuhold was in 1972. Since then it has been a very exciting walk together, not always easy but in the end certainly fruitful. So, looking at IPSI's history and the current situation I think it is fair to say that Erich Neuhold, now retiring, can be proud of what he and his team hand over to his successor. I hope that Erich Neuhold can observe, now from a distance, "his IPSI" – continuously changing and growing in success. I wish he and his family a healthy and fascinating future.

First Introduction, Laudatio, and Personal Note

M. Tamer Özsu

University of Waterloo, School of Computer Science,
Waterloo, Ontario, Canada N2L 3G1
tozsu@db.uwaterloo.ca

I was honored and very pleased to be asked to write an introduction to this volume dedicated to Erich J. Neuhold. I don't quite recall when I first met Erich, which, in itself, is an indication that it was a long time ago. We had overlapping research interests – distributed data management, multimedia systems, heterogeneous databases – so it was natural that we would have a professional relationship. Erich, as the Director of the Fraunhofer Institute (when I first met him this was the German National Research Center for Information Technology, GMD) for Integrated Publication and Information Systems (IPSI) in Darmstadt, Germany, and as the Professor of Computer Science, Publication and Information Systems, at the Technical University of Darmstadt, has created very strong research groups that are conducting state-of-the-art research on a number of fronts. Particularly during the 1990s, IPSI was a place that I visited frequently, including a sabbatical leave in 1998. I have often referred to it as my European base of operations. Erich successfully managed the various technical and organizational transitions that IPSI went through over the years, the last one being the major transition from GMD to Fraunhofer. He, along with his colleagues at IPSI, deserves significant credit for building a world-class research organization.

Erich Neuhold's research over his career has been wide ranging and has covered a variety of topics. His past and present research interests that overlap with mine involve heterogeneous multimedia database systems, distributed systems (mainly peer-to-peer and Grid-based environments), and Web-based distributed information technologies. This volume contains a number of parts that focus on these topics.

The first part in this book is dedicated to issues related to Advanced Technologies for Adaptive Information Management Systems and contains a number of contributions by Erich Neuhold's former PhD students. Thomas Risse, who is currently a division head at Fraunhofer IPSI, reports his latest research in "An Overview on Automatic Capacity Planning". The issue that is addressed is finding a configuration of a distributed system that satisfies performance goals. This is a well-known complex search problem that involves many design parameters, such as hardware selection, job distribution and process configuration. Thomas describes the new approaches that he has been following to address this problem. In the second contribution, "Overview on Decentralized Establishment of Multi-lateral Collaborations", Andreas Wombacher describes service-oriented architectures that facilitate loosely coupled collaborations, which are established in a decentralized way. One challenge for such collaborations is to guarantee consistency, that is, fulfilment of all constraints of individual services and deadlock-freeness. The paper introduces a decentralized approach to consistency checking, which utilizes only bilateral views of the

collaboration. “Dynamic Maintenance of an Integrated Schema” by Regina Motz considers the importance of schema integration methodologies as well as the growing use of cooperative engineering and the relevance of dynamic maintenance of an integrated schema in a federated database. A key characteristic of a federation is the autonomy of its component databases: their instances and schemas may evolve independently, requiring dynamic maintenance of an integrated schema. Regina’s work presents a methodology to propagate structural and semantic modifications that occur in the local schemas of a federated database to the integrated schema. In the final paper of Part 1, entitled “Efficient Evaluation of Nearest-Neighbor Queries in Content-Addressable Networks”, Klemens Böhm, a former PhD student of Erich Neuhold and currently a Professor of Computer Science at the University of Karlsruhe, and Erik Buchmann discuss top-k query processing within peer-to-peer (P2P) systems organized as content-addressable networks (CAN). CANs manage huge sets of (key, value)-pairs and cope with very high workloads while following the P2P paradigm in order to build scalable, distributed data structures on top of the Internet. As outlined in this paper, CANs are designed to drive Internet-scale applications like distributed search engines, multimedia retrieval systems and more. In such scenarios, the introduced top-k query model is very applicable, as the user specifies an objective and the engine responds with a set of the most likely query results.

The second part in this volume is dedicated to issues related to Semantic Web Drivers for Advanced Information Management. In the first contribution, Rudi Studer, a former PhD student of Erich Neuhold and now a Professor of Computer Science at the University of Karlsruhe, and his co-authors Jens Hartmann, Nenad Stojanovic and Lars Schmidt-Thieme focus on Semantic Web mining in a paper entitled “Ontology-Based Query Refinement for Semantic Portals”. Semantic Web mining combines two fast-developing research areas: the Semantic Web and Web mining. On the one hand the authors suggest that this connection can be exploited to improve Web mining methods, using new formal semantics that utilize Web mining results, while on the other hand they build the Semantic Web. The authors also present the SemiPort project, in which methods and tools for semantic information portals are currently developed. In the paper that follows, Matthias Hemmje, another former PhD student of Erich Neuhold and now a Professor of Computer Science at the Distance University of Hagen, and Erich Neuhold’s former PhD student Zhanzi Qiu discuss the issues in preparing legacy Web sites for querying in searches in their paper “Towards Supporting Annotation for Existing Web Pages Enabling Hyperstructure-Based Searching”. They discuss how existing Web pages can be tagged with additional structural and semantic information that can be represented with new Web standards and can be applied in advanced hyperstructure-based Web search methods. This paper is followed by Tom Baker’s fine contribution, entitled “Maintaining Dublin Core as a Semantic Web Vocabulary”, where he describes the Dublin Core Metadata Initiative (DCMI) to maintain a vocabulary of several dozen metadata terms, most notably the 15-element Dublin Core. These metadata terms, along with historical versions of the terms, are identified with URIs, documented on Web pages and in formal schemas, indexed in registries, and cited in application profiles. Tom has been collaborating with researchers at Fraunhofer IPSI in various projects related to such metadata issues.

In addition to his research contributions, it is important to note Erich’s technical contributions more broadly. His original appointment as Director of the then GMD

IPSI (now Fraunhofer IPSI) dates back to 1986. At the same time, he was appointed as the Professor of Computer Science, Chair for Publication and Information Systems, at the Technical University of Darmstadt. Prior to that, he served in various academic positions. He was a Professor of Computer Science, Chair of Applied Informatics and Systems Analysis, at the Technical University of Vienna between 1984 and 1986, and a Professor of Computer Science, Chair of Application Software, at the University of Stuttgart from 1972 to 1983. His research focus was always on information systems and databases, but with wide-ranging approaches. During his academic career, he has been the principal supervisor of 38 PhD students and over 250 master's students.

Erich has held a number of industrial research and development positions as well. While heading the Information Management Laboratory and later the Systems Software Laboratory of Hewlett-Packard Research Laboratories, Palo Alto (1983–1984), Erich took part in and managed research work and development efforts on distributed operating systems, based both on UNIX and new advanced operating system concepts. Communication systems and (distributed) databases were extended to handle multimedia, generalized data types (objects) that were expected to arise in the future office and engineering environments and also in knowledge-based systems. Knowledge-based techniques were investigated in relation to all these areas to determine their possible future impact and to start advanced activities for their incorporation.

During his early times as a research scientist at IBM Corporation (1963–1972) in the US as well as at the IBM laboratory in Vienna, Erich's research work, amongst other aspects, was concerned with synchronization strategies for data manipulation in database systems. Especially during his time at IBM Program Products and Advanced Technology in New York he actively participated in research, design and implementation of integrated databases, database management and information retrieval systems. He contributed to the design and implementation of a family of computational and query languages oriented toward non-professional users.

Finally, let me comment on his service to the scientific community. Erich deserves special thanks with respect to serving as chair/co-chair, organizer/co-organizer, member of program committees, and general supporter in different functions for numerous conferences and other international scientific events in the field. In particular, it is important to note his chairmanship of the International Federation for Information Processing (IFIP) Working Group (WG) 2.6, "Data Bases", and WG 2.2, "Formal Description of Programming Concepts", his Vice-Chairmanship of the IFIP Technical Committee "Information Systems" (TC-8), his membership of the board of the Very Large Data Base (VLDB) Endowment, and his co-presidency of the International Foundation on Cooperative Information Systems (IFCIS). He has also been on the boards of various other committees. Erich has been very active in organizing a large number of conferences – too many to list individually. He serves as the Editor-in-Chief of the *Journal on Digital Libraries*, and he is on the editorial boards of many other journals.

Clearly, a very productive professional life is about to have a phase shift. I do not for a moment expect Erich to take a rest; in fact, I expect to see him and hear him at technical events and meetings for many years to come.

Second Introduction, Laudatio, and Personal Note

Keith van Rijsbergen

University of Glasgow, Department of Computing Science,
17 Lilybank Gardens, Glasgow, G12 8RZ, UK
keith@dcs.gla.ac.uk

Following Tamer Öszus's introduction to the first two parts of this book it is my pleasure to introduce the fourth part of this book. As in the other cases, it reflects Erich Neuhold's research interests, this time in the research areas that are close to my own areas of research: digital libraries and, of course, information retrieval.

The fourth part, aiming to move "From Digital Libraries to Intelligent Knowledge Environments", starts out with a reflection by Edward Fox, Marcos André Gonçalves and Rao Shen on the role of digital libraries in making this move. He illustrates how suitable knowledge environments can be more easily prepared when designed and built on the basis of a variety of digital library metamodels (including those for education, computing and archaeology) thereby resulting in more usable and useful digital library applications. This discussion is followed by some notes by Rudi Schmiede on scientific work and the usage of digital scientific information – its structures, discrepancies, tendencies and strategies. His article discusses changes in scientific work (academic and applied) associated with the new potential benefits accruing from the use of digital library technology, but also the increased coercive powers arising from such use. The background of his interesting observations is discussed in the context of a scientific sphere, in terms of the contents, the quantity and the quality of supply in scientific IT systems, taking into account the user side in their communities of practice, and the technological and organizational basis of scientific information.

Ulrich Thiel and his colleagues at the Fraunhofer IPSI introduce the concept of "Queries in Context" supporting access to digitized historic documents in a so-called Collaboratory, where appropriate indexing and retrieval mechanisms are developed to give users adequate access to material to support their collaborative work. The part continues with Richard Furuta's contribution about the separation of concerns in hypertext where articulation points are expected to increase flexibility. Richard has explored a family of models in which hypertext is modelled by an automaton structure rather than a graph structure. He discusses further distinctions, other than the structure/presentation, that his work provides to interactive documents leading to so-called "articulation points" to support investigations into novel and flexible implementations of hypertext/hypermedia. The fifth paper is a contribution by Karl Aberer, a former division head of Erich Neuhold's Fraunhofer IPSI, now a Professor of Computer Science at EPFL in Geneva, together with Jie Wu. In their paper "Towards a Common Framework for Peer-to-Peer Web Retrieval", they first review existing studies about the algorithmic feasibility of realizing peer-to-peer Web search using text- and link-based retrieval methods. From their perspective realizing peer-to-peer Web retrieval also requires a common framework that enables interoperability of

peers using different peer-to-peer search methods. Therefore they introduce a common framework consisting of an architecture for peer-to-peer information retrieval and a logical framework for distributed ranking computation.

Carol Peters of CNR in Pisa contributed a paper on the comparative evaluation of cross-language information retrieval systems. DELOS, the Network of Excellence on Digital Libraries, has supported the launch of a Cross-language Evaluation Forum (CLEF), with the objective of promoting cross-language information retrieval system development, providing the research community with an infrastructure for testing and evaluating systems operating in multilingual contexts and a common platform for the comparison of methodologies and results. In her paper, Carol outlines the various activities initiated by CLEF over the years in order to meet the emerging needs of the application communities, and traces the impact of these activities on advances in system development.

The final two contributions of this chapter are about providing more intelligent functions within information environments by the application of machine learning techniques. First, in "Personalization for the Web: Learning User Preferences from Text", Giovanni Semeraro of the University of Bari and his colleagues Marco Degemmis and Pasquale Lops deal with the application of supervised machine learning methods for user profiling and content-based information filtering. The paper presents a new method, based on the classical Rocchio algorithm for text categorization, which is able to discover user preferences from the analysis of textual data. Secondly, Thomas Hofmann and Justin Basilico apply collaborative machine learning by proposing a collaborative machine learning framework for exploiting interuser similarities. More specifically, they present a kernel-based learning architecture that generalizes the well-known Support Vector Machine learning approach by enriching content descriptors with interuser correlations.

Besides introducing the fourth part, it is also my pleasure to pay tribute to Erich Neuhold for enthusiastically supporting the publication of several scientific journal series in a variety of responsible functions. He is the former Editor-in-Chief and now, Co-editor-in-Chief of the Journal on Digital Libraries, and a member of the Editorial Advisory Board for ACM Transactions on Database Systems; Information Systems, Pergamon Press; Journal of Information Processing, Japan; IEEE Computing Futures, USA; and Computers & Graphics, Pergamon Press. Finally, he is a member of the Editorial Board for Decision Support Systems – The International Journal, North Holland; Fifth Generation Computer Systems, North Holland; The International Journal of Systems Integration, Springer; Data & Knowledge Engineering, Elsevier Science; Multimedia Tools and Applications – An International Journal, Springer; J.UCS – The Journal of Universal Computer Science, Springer; as well as for the International Journal on Digital Libraries, Springer,

With respect to supporting conferences and other events of the scientific communities in the field of information retrieval, information systems, information science, and digital libraries, over and above those that Tamer has already mentioned in his laudation from the database research communities point of view, Erich has served the ACM SIGIR International Conference on Research and Development in Information Retrieval, the Internationales Symposium für Informationswissenschaft, Digital Libraries (DL), the Forum on Research and Technology Advances in Digital Libraries

(ADL), Advances in Information Systems (ADVIS), the European Conference on Digital Libraries (ECDL), the Joint Conference on Digital Libraries (JCDL), the International Workshop on Digital Libraries (DLib2001), the DELOS Workshop of the International Conference on Digital Libraries (ICDL), as well as Document Image Analysis for Libraries (DIAL).

My own interaction with Erich started in the early 1990s. We began working together on an ESPRIT project IDOMENEUS (1992–1995). My lasting memory of Erich is when he took over as chair of the network and saved the day. It once again illustrated the American dictum: “When the going gets tough, the tough get going”. The network ran to completion and delivered results of which all of us were rightfully proud. Soon after the beginning of that network I was appointed to the position of Beirat of the German National Research Center for Information Technology (GMD) (1993–1996); one of my tasks was to chair a review of the Institute for Integrated Publication and Information Systems in Darmstadt (IPSI), of which Erich had been director since 1986. It is no secret that if the review had gone badly the consequences could have been dire. Erich, in customary fashion, marshalled the troops and oversaw a response to the external review that was exemplary. Once again his courage and stamina saw to it that IPSI continued to flourish.

There is no doubt in my mind that Erich will continue to fight for academic excellence. We all owe him a great debt for the extent of his contributions to our field. Personally, I learnt from him that one can, with one’s friends, overcome bureaucratic obstacles by simply keeping one’s eye on the main goal: success in research. I speak for all when I wish him success in his new ventures.

December 2004

Keith van Rijsbergen

Table of Contents

Advanced Technologies for Adaptive Information Management Systems

An Overview on Automatic Capacity Planning <i>Thomas Risse (Fraunhofer IPSI)</i>	1
Overview on Decentralized Establishment of Multi-lateral Collaborations <i>Andreas Wombacher (University of Twente)</i>	11
Dynamic Maintenance of an Integrated Schema <i>Regina Motz (University of Montevideo)</i>	21
Efficient Evaluation of Nearest-Neighbor Queries in Content-Addressable Networks <i>Erik Buchmann (University of Magdeburg) and Klemens Böhm (University of Karlsruhe)</i>	31

Semantic Web Drivers for Advanced Information Management

Ontology-Based Query Refinement for Semantic Portals <i>Jens Hartmann, Nenad Stojanovic, Rudi Studer, and Lars Schmidt-Thieme (University of Karlsruhe)</i>	41
Towards Supporting Annotation for Existing Web Pages Enabling Hyperstructure-Based Searching <i>Zhanzi Qiu and Matthias Hemmje (FernUniversität Hagen)</i>	51
Maintaining Dublin Core as a Semantic Web Vocabulary <i>Thomas Baker (Fraunhofer-Gesellschaft)</i>	61

Securing Dynamic Media Content Integration and Communication

A Peer-to-Peer Content Distribution Network <i>Oliver Heckmann, Nicolas Liebau, Vasilios Darlagiannis, Axel Bock, Andreas Mauthe, and Ralf Steinmetz (TU Darmstadt)</i>	69
Secure Production of Digital Media <i>Martin Steinebach (Fraunhofer IPSI) and Jana Dittmann (University of Magdeburg)</i>	79

Data Communication Between the German NBC Reconnaissance
Vehicle and Its Control Center Unit
Andreas Meissner and Wolfgang Schönfeld (Fraunhofer IPSI) 87

**From Digital Libraries to Intelligent Knowledge
Environments**

The Role of Digital Libraries in Moving Towards Knowledge
Environments
*Edward A. Fox, Marcos André Gonçalves, and Rao Shen
(Virginia Tech)* 96

Scientific Work and the Usage of Digital Scientific Information –
Some Notes on Structures, Discrepancies, Tendencies, and Strategies
Rudi Schmiede (TU Darmstadt) 107

Queries in Context: Access to Digitized Historic Documents in a
Collaboratory for the Humanities
*Ulrich Thiel, Holger Brocks, Andrea Dirsch-Weigand, André Everts,
Ingo Frommholz, and Adelheit Stein (Fraunhofer IPSI)* 117

Separation of Concerns in Hypertext: Articulation Points That
Increase Flexibility
Richard Furuta (Texas A&M University) 128

Towards a Common Framework for Peer-to-Peer Web Retrieval
Karl Aberer and Jie Wu (EPFL) 138

Comparative Evaluation of Cross-language Information Retrieval
Systems
Carol Peters (ISTI-CNR) 152

Personalization for the Web: Learning User Preferences from Text
*Giovanni Semeraro, Pasquale Lops, and Marco Degemmis
(University of Bari)* 162

Collaborative Machine Learning
*Thomas Hofmann (Fraunhofer IPSI) and Justin Basilico
(Brown University)* 173

**Visualization – Key to External Cognition in Virtual
Information Environments**

Visualization in Digital Libraries
*Enrico Bertini, Tiziana Catarci, Lucia Di Bello, and Stephen
Kimani (University of Rome)* 183

Modelling Interactive, Three-Dimensional Information Visualizations <i>Gerald Jaeschke (Fraunhofer IPSI), Piklu Gupta (TU Darmstadt), and Matthias Hemmje (FernUniversität Hagen)</i>	197
A Knowledge Integration Framework for Information Visualization <i>Jinwook Seo and Ben Shneiderman (University of Maryland)</i>	207
Visualizing Association Rules in a Framework for Visual Data Mining <i>Paolo Buono and Maria Francesca Costabile (University of Bari)</i> ...	221
From Human Computer Interaction to Human Artifact Interaction	
From Human-Computer Interaction to Human-Artefact Interaction: Interaction Design for Smart Environments <i>Norbert A. Streitz (Fraunhofer IPSI)</i>	232
Cooperation in Ubiquitous Computing: An Extended View on Sharing <i>Peter Tandler and Laura Dietz (Fraunhofer IPSI)</i>	241
A Metaphor and User Interface for Managing Access Permissions in Shared Workspace Systems <i>Till Schümmer, Jörg Haake, and Anja Haake (FernUniversität Hagen)</i>	251
Ambient Intelligence: Towards Smart Appliance Ensembles <i>José L. Encarnação and Thomas Kirste (Fraunhofer IGD)</i>	261
Application Domains for Virtual Information and Knowledge Environments	
Enterprise Information Integration – A Semantic Approach <i>Thomas Kamps, Richard Stenzel, Libo Chen, and Lothar Rostek (Fraunhofer IPSI)</i>	271
Ontology-Based Project Management for Acceleration of Innovation Projects <i>Hans-Jörg Bullinger (Fraunhofer-Gesellschaft), Joachim Warschat, Oliver Schumacher, Alexander Slama, and Peter Ohlhausen (Fraunhofer IAO)</i>	280
Understanding and Tailoring Your Scientific Information Environment: A Context-Oriented View on E-Science Support <i>Claudia Niederée, Avaré Stewart, Claudio Muscogiuri (Fraunhofer IPSI), Matthias Hemmje (FernUniversität Hagen), and Thomas Risse (Fraunhofer IPSI)</i>	289

TV Scout: Lowering the Entry Barrier to Personalized TV Program
Recommendation
 Patrick Baudisch (Xerox PARC) and Lars Brueckner
 (TU Darmstadt) 299

Intelligent Home-Enjoying Computing Anywhere
 Junzhong Gu (ECNU-MMIT) 310

Author Index 321