ISBN 0926-5473

IFIP Transactions A-25

INTEROPERABLE DATABASE SYSTEMS (DS-5)

D.K. HSIAO
E.J. NEUHOLD
R. SACKS-DAVIS
Editors

IFIP

NORTH-HOLLAND

INTEROPERABLE DATABASE SYSTEMS (DS-5)

IFIP Transactions A: Computer Science and Technology

International Federation for Information Processing

Technical Committees:

Software: Theory and Practice (TC2)
Education (TC3)
System Modelling and Optimization (TC7)
Information Systems (TC8)
Relationship between Computers and Society (TC9)
Computer Systems Technology (TC10)
Security and Protection in Information Processing Systems (TC11)
Artificial Intelligence (TC12)
Human-Computer Interaction (TC13)
Foundations of Computer Science (SG14)

IFIP Transactions Editorial Policy Board

The IFIP Transactions Editorial Policy Board is responsible for the overall scientific quality of the IFIP Transactions through a stringent review and selection process.

IFIP Transactions Abstracted/Indexed in:
INSPEC Information Services

A-25

INTEROPERABLE DATABASE SYSTEMS (DS-5)

Proceedings of the IFIP WG2.6 Database Semantics Conference on
Interoperable Database Systems (DS-5)
Lorne, Victoria, Australia, 16-20 November, 1992

Edited by

DAVID K. HSIAO
US Naval Postgraduate School
Monterey, CA, USA

ERICH J. NEUHOLD
GMD-IPSI
Darmstadt, Germany

RON SACKS-DAVIS
CITRI
Carlton, Victoria, Australia

1993

NORTH-HOLLAND
AMSTERDAM • LONDON • NEW YORK • TOKYO

ELSEVIER SCIENCE PUBLISHERS B.V.
Sara Burgerhartstraat 25
P.O. Box 211, 1000 AE Amsterdam, The Netherlands

Keywords are chosen from the ACM Computing Reviews Classification System, ©1991, with permission.
Details of the full classification system are available from
ACM, 11 West 42nd St., New York, NY 10036, USA.

ISBN: 0 444 89879 4
ISSN: 0926-5473

This book is printed on acid-free paper.

Printed in The Netherlands

Preface

The proliferation of databases within organizations have made it imperative to allow effective sharing of information from these disparate database systems. In addition, it is desirable that the individual systems must maintain a certain degree of autonomy over their data in order to continue to provide for their existing applications and to support controlled access to their information. Thus it becomes necessary to develop new techniques and build new functionality to interoperate these autonomous database systems and to integrate them into an overall information system. Research into interoperable database systems has advanced substantially over recent years in response to this need.

These considerations led us to organize an International Workshop on the semantics of Interoperable Database Systems. The Workshop is the fifth in a series of workshops on Database Semantics sponsored by the International Federation of Information Processing (IFIP) Working Group 2.6 (Databases).

In response to the Call for Papers, we received approximately 50 full and short papers which were reviewed by members of the Program Committee. Of these papers, 15 full were accepted for regular presentation and 10 were accepted for short presentations at the workshop. The Program Committee also sought 3 invited papers. This volume consists of the 3 invited papers and the 15 regular papers accepted for the workshop.

The papers presented in this volume cover a wide spectrum of both theoretical and pragmatic issues related to the semantics of interoperable database systems. Topics covered include techniques to support the translation between database schema and between database languages; object oriented frameworks for supporting interoperability of heterogeneous databases, knowledge base integration and techniques for overcoming schematic discrepancies in interoperable databases. In addition, there are papers addressing issues of security, transaction processing, data modelling, and object identification in interoperable database systems.

We received excellent support from IFIP Working Group 2.6, the Collaborate Information Technology Research Institute of the University of Melbourne and RMIT (CITRI), the Australian Computer Society, and the US Naval Postgraduate School which made available the facilities for the PC meeting. The workshop was held at a very nice place, providing an inspiring setting for lively discussions far beyond the actual sessions. Our thanks go specifically to John Zeleznikow for his invaluable help in organizing the workshop, and to Melinda Lidstrom, CITRI, for her perfect local arrangements. We would also like to thank all members of the program committee for ensuring the high quality of the proceedings and the authors of the papers presented in these proceedings. In addition, our thanks go to Gisela Fischer, GMD, for her help and support.

We believe that this book will represent a valuable collective contribution to research and development of interoperable database systems.

David K. Hsiao,
Computer Science Department, US Naval Postgraduate School

Erich J. Neuhold,
GMD-Integrated Publication and Information Systems Institute

Ron Sacks-Davis,
CITRI (Collaborative Information Technology Research Institute of the University of Melbourne and RMIT)

This page intentionally left blank

CONTENTS

SPONSORS

IFIP
(International Federation of Information Processing)
Working Group 2.6 (Databases)

CITRI
(Collaborative Information Technology
Research Institute)
and the Key Centre for Knowledge Based Systems,
RMIT and The University of Melbourne

Australian Computer Society

ADDITIONAL SUPPORT:

La Trobe University

ORGANISING BODIES

PROGRAM CO-CHAIRMEN:

David K. Hsiao
Computer Science Department,
US Naval Postgraduate School
Monterey, CA 93943,
USA

Erich J. Neuhold
GMD-Integrated
Publication and Information Systems
Institute
Dolivostr. 15
P.O. Box 10 43 26
D-6100 Darmstadt
Germany

Ron Sacks-Davis
CITRI (Collaborative Information Technology Research
RMIT and The University of Melbourne
723 Swanston Street,
Carlton, Victoria, 3053

TUTORIAL CHAIRMAN:

Robert Meersman
INFOLAB, University of Tilburg,
The Netherlands

GENERAL CHAIRMAN:

John Zeleznikow
Database Research Laboratory, ACRI
La Trobe University, Bundoora,
Victoria, Australia, 3083

LOCAL ARRANGEMENTS:

Melinda Lidstrom
CITRI (Collaborative Information
Technology Research Institute)
723 Swanston Street,
Carlton, Victoia, 3053

PROGRAM COMMITTEE:

P. M. G. Apers
Computer Science Dept
University of Twente
P.O. Box 217
NL-7500 AE Enschede
The Netherlands

S. Ceri
Dipartimento di Ellettronica
Politencnico di Milano
via Ponzio, 34/35
I-20133 Milano
Italy

U. Dayal
DEC
One Kendall Square
Building 700
Cambridge, MA 02139
USA

C. Delobel
INRIA
Domaine de Voluceau
B P 105
F-78153 Le Chesnay Cedex
France

H. D. Ehrich
Institut fuer Programmiersprachen
& Informationssysteme
Abteilung Datenbanken
Technische Universitaet Braunschweig
Postfach 3329
Gauszstr. 12
D-3300 Braunschweig
Germany

K. Dittrich
Universitaet Zuerich
Institut fuer Informatik
Winterthurer Strasse 190
CH-8057 Zuerich
Switzerland

G. Gardarin
MASI Institut Blaise Pascal
Universite P. et M. Curie-CNRS
45, avenue des Etats - Unis
F-78000 Versailles
France

A. K. Elmagarmid
Computer Science Department
Purdue University
West Lafayette, IN 49707
USA

Y. Kambayashi
IMEL, Faculty of Engineering
Kyoto University
Sakyo, Kyoto 606-01
Japan

W. Kent
Hewlett-Packard
1501 Page Mill Road
Palos Altos, CA
USA

W. Klas
GMD-Integrated Publication and
Information Systems Institute
Dolivostr. 15
P.O. Box 10 43 26
D-6100 Darmstadt
Germany

M. Kitsuregawa
University of Tokyo
7-22-1 Roppongi,
Minato-ku
Tokyo 106
Japan

Q. Li
Dept of Computer Science
Hong Kong University of Science
and Technology
Clear Water Bay
Kowloon
Hong Kong

W. Litwin
Department of Computer Science
Standford University
Stanford, CA 94305
USA

F. Lochovsky
Computer Science Department
Hong Kong University of Science
and Technology
Clear Water Bay
Hong Kong

V. Lum
Systems Engineering Department
Chinese University of Hong Kong
Pi Chiu Building
Santin
Hong Kong

D. McLeod
Computer Science Department
University of Southern California
Los Angeles, CA 90087
USA

F. Manola
GTE Laboratories
40 Sylvan Road, MS 62
Waltham, MA 02254
USA

R. Meersman
Tilburg Universiteit
Infolab
P.O. Box 90 153
Hogeschoollaan 225
NL-5000 LE Tilburg
The Netherlands

Y. Masunaga
Dept. of Library and Info Science
University of Library and
Information Science
1-2 Kasuga
Tsukuda-shi, Ibaraki, 305
Japan

B. C. Ooi
Department of Information Systems
Information Systems Institute and
Computer Science
National University of Singapore
Kent Ridge, 0511,
Singapore

M. Papazoglou
School of Information Systems
Queensland University of Technology
GPO Box 2434
Brisbane, Queensland, 4001
Australia

M. Orlowska
Department of Computer Science
University of Queensland
St. Lucia, Queensland, 4067
Australia

K. Ramamohanarao
Department of Computer Science
The University of Melbourne
Parkville, Victoria, 3052
Australia

M. E. Rusinkiewicz
Computer Science Department
University of Houston
Houston, TX 77204
USA

R. Sacks-Davis
CITRI (Collaborative Information
Technology Research Institute)
723 Swanston Street
Carlton Vic 3053
Australia

H-J. Schek
Department of Computer Science
Swiss Federal Institute of Technology
ETH-Zentrum
CH-8092 Zuerich
Switzerland

P. Scheuermann
School of Engineering and Applied Sciences
Northwestern University
2145 Sheridan Road
Evanston, IL 60208
USA

A. P. Sheth
Bellcore
RRC 1j-210, 444 Hoes Lane
Piscataway, NJ 08854
USA

A. Silberschatz
Computer Science Department
University of Texas at Austin
Austin, TX 78712-1184
USA

S. Spaccapietra
Computer Science Department
Swiss Federal Institute of Technology
CH-1000 Lausanne
Switzerland

R. Studer
Institut fuer Angewandte Informatik
und formale Beschreibungsverfahren
Universitaet Karlsruhe
Kollegiengebaeube am
Ehrenhof
Postfach 6980
D-7500 Karlsruhe 1
Germany

A. M. Tjoa
Institut fuer Statistik und Informatik
Universitaet Wien
Liebiggasse 4
A-1010 Vienna
Austria

K-Y. Whang
Computer Science Department
Korea Advanced Institute
of Science and Technology
P.O. Box 150
Cheng-Ryang Ui, Seoul
Korea

C. T. Wu
Code CS, Computer Science Dept
Naval Postgraduate School
Monterey, CA 93943
USA

C. Yu
Department of EE and CS
University of Illinois at Chicago
Chicago, IL 60680
USA

ADDITIONAL REFEREES:

H. Dieudonne
P. Fankhauser
T. Hartmann
K. Jambor
H Y.J. Kim
M. McLoughlin
J. S. Park
F. Saltor
W.K. Whang
G. Dong
M. Gogolla
J. Hiller
B. Jiang
M.L. Lee
P. Muth
J-T. Park
S. Suzuki
Y. Zhang
L. Dunn
M. Hartig
E.K. Hong
M. Kaul
H. Lu
K. Neumann
Y. C. Park
K. Wang

This page intentionally left blank

Interoperable Database Systems (DS-5) (A-25)
D.K. Hsiao, E.J. Neuhold and R. Sacks-Davis (Editors)
Elsevier Science Publishers B.V. (North-Holland)

The promise of distributed computing and the challenges of legacy information systems[1]

Michael L. Brodie

Intelligent Database Systems
GTE Laboratories Incorporated, 40 Sylvan Road, Waltham, MA 02254
brodie@gte.com

Abstract

The imminent combination of computing and telecommunications is leading to a compelling vision of world-wide computing. The vision is described in terms of next generation computing architectures, called Enterprise Information Architectures, and next generation information systems, called Intelligent and Cooperative Information Systems. Basic research directions and challenges are described as generalizations of database concepts, including semantic aspects of interoperable information systems. No matter how compelling and potentially valuable the vision may be, it is of little use until the legacy problem is solved. The problem of legacy information systems migration is described, in the context of distributed computing, and is illustrated with lessons learned from actual case studies. The basic research directions and challenges are recast in the light of actual legacy information systems. Recommendations for both realizing the vision and meeting the challenges are given, including the search for the elusive Killer Application and one fundamental challenge for future information systems technology.

Keyword Codes: H.0; H.1.1; H.4.0
Keywords: Information Systems, General; Systems and Information Theory, Information Systems Applications, General

1. WORLD-WIDE COMPUTING

My professional goal is to contribute to making the world a better place by providing solutions to significant, practical problems (see Appendix). As a computer science researcher, this means that I want to produce the highest quality research and technology that is ultimately applicable to real problems so that the results are consistent with my beliefs. In this regard, I have high hopes and expectations for the potential benefits of world-wide computing. The vision is that problems or questions posed by one or more agents (e.g., humans or computer) can be solved as automatically and transparently as possible. Automatically means that the necessary computing resources [e.g., programs, information bases, information systems (ISs)] are identified and caused to interact cooperatively to effectively and efficiently solve the problem. Transparency means that all unnecessary details are not seen by the agent (e.g., locations and nature of the participating resources).

In this section, I describe a world-wide computing vision in terms of cooperation amongst ISs augmented by a telecommunications vision that provides communication on a scale previously unthinkable by computer scientists.

---

[1]An earlier version of this paper appeared in P.M.D. Gray and R.J. Lucas (eds.) *Advanced Database Systems: Proceedings of the 10th British National Conference on Databases*, Springer-Verlag, New York/Heidelburg, 1992.

1.1. The Vision

The vision of distributed computing is compelling. It says that soon the dominant computing paradigm will involve large numbers of heterogeneous, intelligent agents distributed over large computer/communication networks. Agents may be humans, humans interacting with computers, humans working with computer support, and computer systems performing tasks without human intervention. Work will be conducted on the network in many forms. Work task definition will be centralized (e.g., a complex engineering task) and decentralized. Tasks will be executed by agents acting autonomously, cooperatively, or collaboratively, depending on the resources required to complete the task (e.g., monitoring many systems of a patient or many stations in a factory). Agents will request and acquire resources (e.g., processing, knowledge, data) without knowing what resources are required, how to acquire them, or how they will be orchestrated to achieve the desired result. A goal of this vision is to be able to use, efficiently and transparently, all computing resources that are available on computers in large computer/communications networks.

1.1.1. Cooperative Work

Computers should support humans and organizations in their natural modes of thinking, playing, and working. Consider how complex activities are in human organizations such as a hospital (Figure 1). Each human agent (e.g., doctor, technician, nurse, receptionist) provides capabilities to cooperatively achieve a goal (e.g., improve the health of a patient). For a doctor to complete an analysis of a patient, the doctor may need the opinion of another doctor, the results of a laboratory test, and personal information about the patient. In general, the analysis is broken into sub-activities and appropriate agents are found for each sub-activity. Each sub-activity is sent to the appropriate agents together with the required information in a form that the agent can use. Cooperating agents complete the sub-activities and return the results in a form that the doctor can use. The doctor then analyzes the results and combines them to complete the analysis, possibly by repeating sub-activities that were not successful or by invoking new sub-activities.

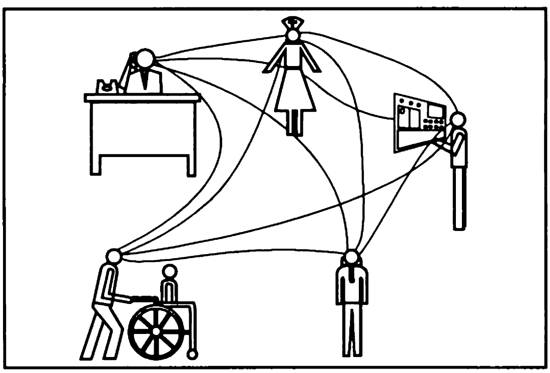

Figure 1. Cooperating agents in medical care.

Such cooperative work requires considerable intelligent interaction among the agents using knowledge of who does what, what information is required, the form in which it is required, scheduling requirements or coordination of tasks, how to locate agents, how to request that sub-activities be done, etc. The cost and quality of products of most human organizations depend on the effectiveness of such cooperation. In hospitals, the quality and cost of health care depend on effectiveness and speed of cooperation. Aspects of the cooperation can be seen as effective parts of the work being done (e.g., doctor's interaction to solve life critical problems), while others may be seen as counterproductive (e.g., converting patient chart information into multiple computer formats for automated analysis steps). The cost and complexity of interactions in a hospital argue for their optimization. What cooperation aspects are effective and should be encouraged, and which should be diminished?

Intuitively, it seems that the distributed computing vision could meet many requirements of cooperative work. The cost of an activity could be reduced by a computing infrastructure that makes appropriate interactions transparent to the agents. Computers could contribute to more productive (e.g., effective and efficient) work by intelligently supporting cooperation. In the next section, I examine forms of intelligence and cooperation that computers might support. I limit my scope to the cooperative work that might be supported by cooperating ISs and the resulting requirements on the computing infrastructure, or systems technology.

1.1.2. Intelligent and Cooperative Information Systems

Intelligent and Cooperative Information Systems (ICISs) are seen as the next generation of IS, 5-10 years in the future [BROD92a]. ICISs are collections of ISs that exhibit forms of cooperation and intelligence. Cooperation is supported by interoperability (the ability to interact effectively to achieve shared goals, e.g., a joint activity). Intelligent refers, in part, to the ability to do this efficiently (i.e., have the system find, acquire, and orchestrate resources in some optimal fashion) and transparently (with the least human effort). The goal is that any computing resource (e.g., data, information, knowledge, function) should be able to transparently and efficiently utilize any other. Although some features of such systems are agreed upon, no one knows the exact nature of these systems. This sub-section illustrates and suggests some initial ideas for ICIS functionality.

Most organizations have developed many application-specific but independent ISs and other computing resources. They soon find that almost all ISs must interact with other ISs or resources, just as the people in their organizations need to interact. Such organizations have vast investments in valuable resources that cannot be used together without great cost. For example, valuable data is bound to applications and is not available to others. There is a growing need for vast numbers of disjoint information/computing resources to be used cooperatively, efficiently, transparently, and easily by human users (e.g., clerks, scientists, engineers, managers). Consider, for example, the different ISs that must interact to support the functions of a hospital (Figure 2). To produce a patient bill, the billing system must obtain information from many hospital ISs (e.g., nursing records, doctors' bills, pharmacy, radiology, lab, ward, food services).

Let's call such an effective combination of systems, a Health Care ICIS. The Health Care ICIS requires access between multiple, heterogeneous, distributed ISs that were independently designed to be used in isolation.

I consider two or more ISs that execute joint tasks to form a larger IS, called a *cooperative IS*. I call an individual IS within a cooperative IS a *component IS*. With various forms of transparency, a cooperative IS can act as, and be considered as, a single IS (e.g., the hospital billing system accesses of multiple ISs should be transparent to the user). A common requirement for component ISs is to maintain autonomy while cooperating within the cooperative IS.

Intelligent features could be added to a cooperative IS. These features require of technology, or provide users with, more *intelligence* than do conventional ISs. Intelligence has a potential role in user interaction between the user and the component ISs to enhance the quality of interaction. Examples of such features include presenting an integrated view of the multiple ISs; explanation; intentional queries; and presenting functionality through graphic, visual, linguistic or other support (e.g., use of icons, templates, graph representations).

Intelligence also plays a role in enhancing IS functionality. Examples include the following:

- Enhanced decision making or reasoning capabilities (e.g., incorporate hospital rules into the Health Care ICIS).
- (Re) Active (e.g., when a new patient is registered, a transaction is triggered that checks the availability of rooms in wards and orders needed supplies).
- Nondeterminism (e.g., give me any one of the possible teams that has two doctors from cardiology, an anesthetist, and three nurses who are not already booked).
- Nondeductive forms of inference (e.g., induction such as learning rules or constraints from databases, reorganizing a schema based on current extensions of different classes, redistributing information based on access patterns; case-based reasoning, where information is structured according to *cases* and new situations are dealt with by finding similar ones in the information).
- Maintaining integrity constraints.
- Introspection: reasoning about meta-knowledge (e.g., a Health Care ICIS component reasoning about what it can and cannot do in the face of a request).

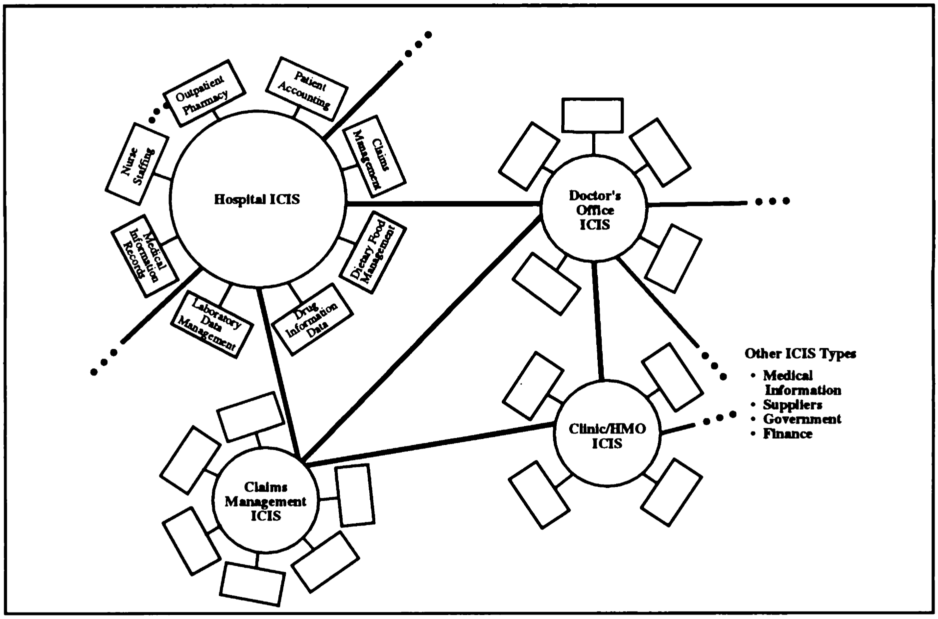

Figure 2. Health care ICIS.

1.1.3. The Global Computer

In a separate universe far away, or so it seems, a vision for the next generation telecommunications technology is taking shape. It intends to permit any information to be communicated anywhere, at any time, in any form, to any agent (human or machine). The key technologies include ubiquitous, broadband, lightning fast intelligent networks enabling