

STATE OF THE ART REPORT

DATABASE~ THE 2nd GENERATION



SERIES 10
NUMBER 7

STATE OF THE ART REPORT

DATABASE~ THE 2nd GENERATION

EDITOR: D J L GRADWELL,



**SERIES 10
NUMBER 7**

THE STATE OF THE ART REPORTS - 1982 SERIES

- 1 Network Architectures
- 2 Programming Technologies
- 3 Business Graphics
- 4 Computer Operations
- 5 Integrated Office Systems
- 6 Supercomputer Systems Technology
- 7 Database - The Second Generation
- 8 Technology Management

Published by Pergamon Infotech Limited,
Maidenhead, Berkshire, England.

Printed by A Wheaton & Company Limited,
Exeter, Devonshire, England.

UDC 681.3

Dewey 658.505

ISBN 008 028570 8

© Pergamon Infotech Limited, 1982

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photographic, or otherwise, without the prior permission of the copyright owner.

DATABASE~ THE 2nd GENERATION

CONTENTS

<u>Editor's foreword</u>	1
--------------------------	---

<u>Publisher's note</u>	4
-------------------------	---

ANALYSIS

1	<u>THE SECOND-GENERATION MARKET-PLACE</u>	5
1:1	THE SLOW ARRIVAL	7
	INTRODUCTION	7
	THE DBMS MARKET IS BOOMING	8
	DBMSs FOR MICROCOMPUTERS	9
	THE ARRIVAL OF COMMERCIAL RELATIONAL DATABASE SYSTEMS	9
	METHODOLOGIES FOR SYSTEMS ANALYSIS AND DESIGN	10
	PUBLICLY AVAILABLE DATABASES	11
	ACCESS TO DATABASES THROUGH VIDEOTEX NETWORKS	12
1:2	DBMS ON MICROCOMPUTERS	17
	THE ARRIVAL	17
	EXAMPLES OF DBMSs FOR MICROS	18
	CARD INDEX SYSTEMS	25
	DIRECTIONS FOR DEVELOPMENT OF SELF-CONTAINED SYSTEMS	26
	AN ALTERNATIVE TO THE DBMS ON THE MICRO ITSELF	27
	WHAT IS A DBMS?	28
1:3	PORTABILITY OF DATABASE APPLICATIONS	29
	INTRODUCTION	29
	INDUSTRY STANDARD OPERATING SYSTEMS	29
	DBMSs AVAILABLE ON SEVERAL TYPES OF HARDWARE	30
	BACK-END DATABASE MACHINES	30
	SPECIAL FACILITIES TO ALLOW PROGRAMS WRITTEN FOR ONE DBMS TO ACCESS ANOTHER	32
	NEW DBMSs THAT ENCOMPASS EARLIER ONES	33
1:4	EXPERIENCE WITH CURRENT DATABASE PRODUCTS	35
	RECORDING EXPERIENCE	35
	USER EXPERIENCE	35
	METHODOLOGIES NEEDED	37
	USE OF A DATA DICTIONARY	40

PHYSICAL DESIGN OF THE DATABASE	42
ORGANISATIONAL STRUCTURES NEEDED	43
COMMON APPLICATION DESIGN PROBLEMS	46
WHAT ARE THE REALLY IMPORTANT ASPECTS OF USING DATABASE TECHNOLOGY?	46
PERFORMANCE MEASUREMENT	48
 2 <u>DATABASE APPLICATION DEVELOPMENT</u>	 59
2:1 DEVELOPMENT ENVIRONMENTS AND DDSs	61
DIRECTIONS IN LANGUAGE DEVELOPMENTS	61
AUTOMATING THE JOBS OF BUILDING APPLICATIONS	62
OBJECTIVES AND SCOPE OF A DDS	64
AT THE CENTRE OF A SET OF APPLICATION DEVELOPMENT TOOLS	69
THE DATA DICTIONARY AS THE HOLDER OF COMPLETE PROGRAM DEFINITIONS	70
USING THE EXTENSIBILITY FEATURE OF DATA DICTIONARIES	72
PRIVACY AND DATA DICTIONARIES	76
DBMS ARCHITECTURE AND DATA DICTIONARIES	77
FUTURE DATA DICTIONARY DEVELOPMENTS	78
2:2 END-USER INTERFACES TO DATABASE SYSTEMS	81
RAPID DEVELOPMENT	81
HOW HAS DATABASE TECHNOLOGY BROUGHT END USERS CLOSER TO THEIR DATA?	82
CURRENT QUERY SYSTEMS	82
WHAT DO WE REQUIRE OF A QUERY LANGUAGE SYSTEM?	83
ACCESS TO DATA IN MORE THAN ONE TYPE OF DATABASE AT THE SAME TIME	93
MULTIBASE	94
MAPPING THE GLOBAL VIEW TO THE VARIOUS LOCAL VIEWS	98
COLOUR GRAPHICS OUTPUT FROM DATABASES	100
VISION OF THE EXECUTIVE'S FUTURE	105
2:3 FOURTH-GENERATION PROGRAMMING LANGUAGES FOR DATABASE SYSTEMS	107
SOLVING THE SOFTWARE CRISIS	107
INTRODUCTION TO SOME SYSTEMS	109
ROLE OF THE DATA DICTIONARY IN FOURTH-GENERATION PROGRAMMING SYSTEMS	110
PROCEDURALITY	114
SPECIFYING ACCESS TO THE DATABASE	116
FACILITIES FOR DIALOGUE DESIGN	120
PROTOTYPING	121
CONCLUSIONS ABOUT FOURTH-GENERATION PROGRAMMING SYSTEMS	123
2:4 DEVELOPMENT IN SYSTEMS ANALYSIS METHODOLOGIES	125
THE NEED FOR NEW METHODOLOGIES	125
DP PRACTITIONERS	127
THE NEED FOR A FORMAL DOCUMENTATION TECHNIQUE	128
THE DEVELOPMENT OF ANALYSIS METHODOLOGIES	130
ENTITY LIFE-CYCLE MODELS	134
COMPARISONS OF METHODOLOGIES AND MODELLING TECHNIQUES	135
CONCLUSIONS ON IMPROVED SYSTEMS ANALYSIS TECHNIQUES	139
2:5 DATABASE DESIGN TOOLS	141

3	<u>DBMS ARCHITECTURE IN THE SECOND GENERATION</u>	151
3:1	CHANGES FROM THE FIRST TO THE SECOND GENERATION	153
	A NEW GENERATION	153
	DIRECTIONS	153
3:2	RELATIONAL DBMSs - DEVELOPMENTS 1981/82	155
	NEW FULL FUNCTIONALITY RELATIONAL SYSTEMS	155
	PROVISION OF TRANSACTION-BASED RECOVERY	158
	CHANGING THE LOGICAL STRUCTURE OF THE DATABASE	160
	CHANGING THE PHYSICAL STRUCTURE OF THE DATABASE	160
	RELATIONAL SYSTEMS AS THE BASIS FOR PERSONAL DATA HANDLING	162
	TOOLS FOR APPLICATION CREATION	164
	WORD PROCESSING LINKED TO A RELATIONAL DATABASE QUERY LANGUAGE	172
	TEXT HANDLING IN RELATIONAL DATABASES	172
	CONCLUSION ON RELATIONAL DBMSs	173
3:3	BACK-END DATABASE MACHINES REVISITED	175
	THE OVERALL ARCHITECTURE OF BACK-END DATABASE MACHINES	178
	REASONS TO BUY A BACK-END DATABASE SYSTEM	180
	THE NEED FOR AN INTEGRATED HARDWARE/SOFTWARE SOLUTION	187
	FIRST STEPS TOWARDS DISTRIBUTED DATABASES	191
	PROBLEM AREAS WITH BACK-END DATABASE MACHINES	194
	FILE SERVERS FOR LANs	196
3:4	OPEN SYSTEMS INTERCONNECTION	199
	FEDERATED DATABASES	199
	ISO AND ECMA WORK ON DRAFT STANDARDS FOR OSI	201
	THE BEGINNINGS OF A PROTOCOL FOR FILE TRANSFER, ACCESS, AND MANAGEMENT	206
	OSI CONCLUSIONS	213
3:5	DBMS STANDARDISATION	215
	WHY DO WE NEED STANDARDS FOR DBMSs?	215
	THE LONG ROAD TO STANDARDS	216
	A POSSIBLE ARCHITECTURE FOR DATABASE STANDARDS?	217
	STANDARDS FOR DDSs	220
	PROBLEM AREAS IN STANDARDISING DATA DICTIONARIES	224
	STANDARDS FOR RELATIONAL DBMSs	226
	A DRAFT ANSI STANDARD FOR NETWORK STRUCTURED DATABASES	227
	THE WORK OF ISO	228
	GERMAN STANDARDISATION ACTIVITIES	229
	STANDARDS - A CONCLUSION	231
3:6	DISTRIBUTED DATABASES	233
	PLANNING A DISTRIBUTED DATABASE	233
	MATCHING INFORMATION SYSTEM DEVELOPMENT WITH COMPANY STRATEGY	237
	SOME DISTRIBUTED DATABASE SYSTEMS	240
	APPROPRIATE PROGRAMMING LANGUAGES	247
3:7	DATA MODELS; NETWORK AND RELATIONAL; UNIFICATION OR DIVERGENCE?	249
	WHY HAS THE DP INDUSTRY HAD SO MUCH TROUBLE WITH DATA MODELLING	249
	ESCAPING FROM THE CONSTRAINTS OF A RECORD-BASED MODEL OF DATA	250
	PRACTICAL PRODUCT DEVELOPMENTS	251
	THE RELATIONAL VERSUS NETWORK DISPUTE	254
	IMPROVING PROGRAMMING PRODUCTIVITY USING THE RELATIONAL MODEL	278

4	<u>CONCLUSION</u>	281
	TOWARDS THE NEXT GENERATION OF DATABASE SOFTWARE	283
	STATE OF FLUX	283
	REQUIREMENTS FOR THE NEXT GENERATION OF DBMS	284
	WHITHER RELATIONAL SYSTEMS?	285
	WHAT IS A DBMS?	285
	INVITED PAPERS	
	IMS - PARTICULAR ISSUES, GENERAL DIRECTIONS	287
	<i>I R Aitken</i>	
	SYSTEM DEVELOPMENT FACILITIES	319
	<i>M Anderson</i>	
	ESTIMATING AND MEASURING DATABASE PERFORMANCE - A PRACTICAL EXAMPLE	331
	<i>C Baker and M J R Keen</i>	
	THE REASONS FOR THE COMMERCIAL SUCCESS OF THE INTEL DATABASE PROCESSOR	347
	<i>W E S Conings</i>	
	RADS - A PROGRAMMING LANGUAGE FOR THE NEXT GENERATION OF DATABASE MANAGEMENT SYSTEMS	358
	<i>H G Cosh</i>	
	CULLINANE'S APPLICATION DEVELOPMENT SYSTEM: STATE OF THE ART PROGRAMMING FOR ON-LINE DATABASES	377
	<i>J de Martino</i>	
	THE ANALYSIS OF INFORMATION NEEDS - SOME PRACTICAL EXAMPLES 1965 TO 1982	395
	<i>H Ellis</i>	
	WHAT MAKES THE ADABAS DATA BASE MACHINE COMMERCIALY SUCCESSFUL?	417
	<i>S D Fitches</i>	
	USING DATABASE TECHNOLOGY - A CASE STUDY	435
	<i>B Francis</i>	
	INTEGRATING DATA DICTIONARY SYSTEMS WITH DBMS	443
	<i>D J L Gradwell</i>	
	DATA DICTIONARY FEATURES FOR THE 1980s	461
	<i>J Hackett and C Warner</i>	
	DISTRIBUTED DATABASES	475
	<i>R H Holland</i>	
	MDBS III: A FULL SCALE DBMS FOR MICROCOMPUTER SYSTEMS	499
	<i>C W Holsapple</i>	
	THE CENTRAL ROLE OF DICTIONARY SYSTEMS IN INFORMATION RESOURCE MANAGEMENT	519
	<i>A Mayne</i>	

THE HIGH-LEVEL END USER	535
<i>R Meersman</i>	
USING DATABASE TECHNOLOGY - A CASE STUDY OF THE COSTS AND BENEFITS	555
<i>P Mercer</i>	
DATABASE MACHINES IN PRACTICE	565
<i>T F Meurer</i>	
CONCEPTUAL MODELLING AS A TOOL FOR APPLICATIONS DEVELOPMENT	583
<i>T W Olle</i>	
USING DATABASE TECHNOLOGY: A CASE STUDY OF THE COSTS AND BENEFITS	593
<i>M J Pratt and H Clough</i>	
THE DESIGN OF AN END-USER QUERY INTERFACE TO A CORPORATE DATABASE	605
<i>R M Tagg</i>	
 BIBLIOGRAPHY AND ANALYSIS REFERENCES	
AN ANNOTATED BIBLIOGRAPHY OF DATABASE - THE 2nd GENERATION	621
<i>D J L Gradwell</i>	
 INDEX	
SUBJECT/CONTRIBUTOR INDEX	651

EDITOR'S FOREWORD

In this Report we review the new generation of database products and ideas. Our goal is to portray an overall picture of the landscape: it is to draw a map: it is to guide you to those topics which you will find most interesting, even fascinating. To do this we show enough of the products, the ideas, the disagreements, even sometimes the characters of the protagonists, to lead you to follow the references and hence to the Annotated Bibliography and on to the original papers.

A Report such as this can only skim the surface of what has become an enormous topic. The references for this Report fill several metres of bookshelves and so we have been selective in the material covered. There are a number of previous State of the Art Reports on database technology or related to it, which we have tried not to duplicate, but rather have drawn from to point out that which is new and interesting.

Thus this Report has a somewhat practical slant. A great deal of it deals not only with academic research but with products and users and the consultants who advise them. We look at the changing market-place with its new microcomputer-based DBMSs; we look at the experience of users of the first generation of database systems. Then we turn to the major problem facing the DP industry, that of application development. We look at the tools that have grown up based upon database technology: data dictionaries, application development systems or fourth-generation programming languages, call them what you will. Then we consider methodologies for systems analysis for database systems; we look at some of the tools for database design.

The architecture of DBMSs is changing. Efficient, complete, relational systems have come to the market; back-end database machines offer the possibility of significant reduction in the cost of employing database software. They have the potential for more efficient hardware unencumbered by current general-purpose operating systems. Then there is the possibility of connecting different databases controlled by different DBMSs. The International Standards Organisation (ISO) is working on standards for Open Systems Interconnection. Other standards for database systems are arriving, but very very slowly. Perhaps the rate of change in database technology makes it impossible

to produce a standard before it is obsolete. For the sake of the users of database systems we hope that this is not the case.

Finally we look at the continuing debate between the proponents of the network data models and the relational data model. We let Adiba give his view of the fate of relational database systems.

We are left with two trains of thought. The first is that the database industry is in a state of flux. A few years ago it had some appearance of stability. Now, however, the rate of change is quickening. What is more, the indications are that this rate of change will accelerate. We see Martin's statement that 60% of new applications will be using database technology by 1983. Even if this is a little optimistic, such a rate of usage will place immense pressures on vendors of database systems. They will have to make their systems highly usable by untrained staff or they will disappear under the support load.

The second train of thought is that we are no longer sure exactly what a DBMS is. When we first decided on the major section headings for this Report, the distinction between relational DBMSs and application development systems such as we discuss in the section on fourth-generation languages was very clear. However, after looking at some of the facilities provided by the relational systems INGRES and ORACLE, we are no longer very clear at all! Thus readers interested in fourth-generation language systems should read the section on relational systems as well.

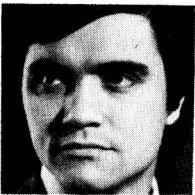
A similar problem arose over the split between query languages and relational systems. Again, anyone interested in query systems should read the section on relational products as well. To add to our confusion the difference between self-contained database systems for end users and application development systems steadily disappeared. The difference between DBMS for microcomputers and application development systems became muddier. It seems that there are no clear distinctions any more between the various software components. We are not sure where this will lead. Will everyone buy a single totally integrated package from their chosen vendor? Will it do everything? We see all the vendors rushing in this direction.

Then we read the NBS proposals for an architecture for DBMS standards, it proposes that a component from one vendor works together with a component from another. Can this be possible?

We have illustrated these trends, concepts, and ideas by referring to a large number of commercial products. By choosing any particular product we do not mean to endorse it as particularly good or particularly bad of its kind. Nor should anyone selecting

a product only consider those mentioned. We have omitted mention of many. This Report is not intended to be a product review, rather, it seeks to show directions; it seeks to explain the many points of view. It will hopefully lead you to a deeper understanding of the impact of database technology on DP in general and to a wish to pursue the areas of debate still further.

David J L Gradwell: Editor



D J L GRADWELL graduated from Cambridge in 1971. He has specialised in database design (both for CODASYL type DBMSs and IBM's IMS) and the design and implementation of data dictionaries. In 1974 he founded the British Computer Society's Data Dictionary Systems Working Party which he chaired for four years. He has been a consultant to many international companies on all aspects of data administration, particularly data and functional analysis, database design and data dictionary evaluation and selection. Mr Gradwell is currently Chief Designer - Usability for ICL Data Dictionary Systems and related products.

PUBLISHER'S NOTE

This Report is divided into three parts:

- 1 ANALYSIS.
- 2 INVITED PAPERS.
- 3 BIBLIOGRAPHY.

The Analysis

- 1 Serves as an introduction to the major technical advances in database systems.
- 2 Provides a balanced analysis of the state of the art in database systems.

The Analysis is constructed by the editor of the Report in the form of a discussion among the chief authorities in this field. The quotations from these authorities are selected by the editor to present a balanced and comprehensive view of the latest developments in database systems. To enable readers to follow up points of interest, the sources of all quotations are given in the Bibliography which follows the Invited Papers.

The following editorial conventions are used throughout the Analysis:

- 1 Material in Courier italic (this typeface) is written by the editor.
- 2 Material in Courier roman (this typeface) is contributed by the person or publication whose name precedes it. The contributor's name is set in Diplomat; a number in the range 100-199 in brackets following the name refers to the original source, as specified in the Bibliography. Numbers within the text from 200 onwards follow the same convention. A name without a reference refers to an Invited Paper published in this Report.
- 3 The quotations in the Analysis are arranged at the discretion of the editor to bring out key issues. Three or four dots within a single quotation indicate that a portion of the original text has been removed by the editor to improve clarity.

The Invited Papers in this State of the Art Report examine various aspects of database systems. If a paper cites references they are given at the end of that paper, numbered in the range 001-099.

The Bibliography is a specially selected compilation of the most important published material on the subject of database systems. Each key item of the literature is reviewed and annotated to assist in selecting the required information.

COPYRIGHT

It should be noted that the copyright for the Report as a whole belongs to Pergamon Infotech Ltd. The copyright for individual contributions belongs to the authors themselves.

1: The second generation market~place

Until recently database technology has been restricted to companies with large machines and large budgets. Now, however, the shape of the market-place is changing: new systems are appearing for smaller computers. In this first section an overview is given of the changes that are occurring. The following three aspects of today's database market-place are examined:

- The exciting prospect of DBMSs for microcomputers
- The perennial problem of the lack of portability of database applications
- The experience of those who have tried in earnest to use database systems.

THE SLOW ARRIVAL

INTRODUCTION

After 10 years of struggle, many failures, many developments more costly than anticipated, the use of database management systems (DBMSs) is becoming the norm. It is no longer the preserve of DP departments who consider themselves at the leading edge and are willing to pay for that privilege. DBMSs are now being sold for the largest of mainframes, the smaller mainframes, the minicomputers and even for the smallest of micros.

In many ways, the industry is showing signs of maturity. Several major database systems have hundreds if not thousands of users. Most of the major systems have all the security, recovery, back-up, concurrency control and other detailed facilities that allow the construction of large on-line databases with many terminals and complex transactions. Analysis and design methodologies that five years ago were considered avant-garde are now treated as the standard way to work.

Yet there seems to be a tremendous conflict between this view of stability and steady, almost boring, progress towards the normality of every day data processing and the view one gets when analysing the changes in the database industry. We see a stunning rate of product development. We see an increasing rate of product development. This Report tries to capture some highlights of the newer products.

We see the term 'database management system' losing its once clear meaning. In the micromarket the term can mean anything from a replacement card index to something a bit better than what we used to call a complete DBMS. In the minicomputer market several DBMSs come complete with a substantial amount of application development software. The difference between a programming language and a DBMS is no longer clear.

Once a DBMS was just a piece of software. Now the vendors of back-end database machines are doing their best to point out the benefits of specialist hardware unencumbered by conventional operating systems.

Meanwhile the research into distributed databases continues and some distributed database products become commercial realities. In this Report we review the many aspects that have become a part of database technology today. We look from the micro to the mainframe; from the end user to the programmer facing a fourth generation of programming languages; from real experience with today's database products to research systems; from the practicality of today's systems to the possibilities for tomorrow.

THE DBMS MARKET IS BOOMING

During the Seventies the number of database users crept up slowly. They were usually very large companies with resources to risk on investing in a new technology. Now, however, it is becoming more of the norm to use database technology rather than not to. This has a major impact on the market-place.

KRASS and WIENER (147): DBMS vendors and users alike agree that the market for the systems is going to take off during the Eighties.

Edward Cherlin, for instance, expects the DBMS market to grow at a compounded 30% right through 1989. Cherlin is a senior analyst at Strategic Inc, the San Jose California market research organisation. He puts the current market at something like \$500 million a year. By the end of the decade, he figures, the total expenditures on software plus hardware (which means database engines, not the mainframes they hang on) will be in the neighbourhood of \$4 billion. He believes that the database machine portion of that total will be \$1.2 billion.

Krass and Wiener list a mere 54 database products in their survey (147). This is a much greater range than even five years ago.

In this Report we shall look at several growth areas for database technology. These include:

- Database systems for microcomputers
- Relational database systems
- Back-end database systems
- Distributed database systems.