

Origins of the Data Base Management System “Bucket of Facts”

Anita Bhatia

*Dronacharya Institute Of Management and Technology, Kurukshetra, Haryana
anitahanda29@gmail.com*

ABSTRACT

The data base concept derives from early military on-line systems, and was not originally associated with the specific technologies of modern data base management systems. While the idea of an integrated data base, or “bucket of facts, ” spread into corporate data processing and management circles during the early 1960s, it was seldom realized in practice. File-processing packages were among the very first distributed as supported products, but only in the late 1960s were they first called “data base management systems, ” in large part through the actions of the Data Base Task Group of the Committee on Data Systems Languages (CODASYL). As the DBMS concept spread, the data base itself was effectively redefined as the informational content of a packaged DBMS.

1. INTRODUCTION

The Data Base Management System (DBMS) is the foundation of almost every modern business information system. Virtually every administrative process in business, science or government relies on a data base. The rise of the Internet has only accelerated this trend – today a flurry of database transactions powers each content update of a major website, literature search, or internet shopping trip. Yet very little research addresses the history of this vital technology, or that of the ideas behind it. We know little about its technical evolution, and still less about how its usage has changed over time.

On a technical level, however, the DBMS evolved from a more humble class of programs known as “file management systems”, created within the unglamorous world of corporate data processing to simplify the creation of programs for routine administration. The data base management system conflated the managerial concept of the data base with the specific technology of the file management system. As this paper shows, in practice the DBMS worked well as a technical system to aid

application programmers, but disappointed as a managerial panacea. Most early DBMS systems were used primarily for routine applications, were not queried directly by managers, and did not support the integration of all corporate data. In addition, while the corporate data base had originally been conceived as a repository of all important managerial information, actual DBMS technology supported only the kind of highly structured regular records with which earlier file management systems had been adept.

2. THE DATA BASE AND THE MANAGEMENT INFORMATION SYSTEM

During the 1970s, when data base management systems were first promoted to corporate managers, they were sold as the technological means by which all of a company's computerized information could be assimilated into a single integrated pool of data. This idea was not, however, a new one. Indeed, its widespread discussion among experts on the managerial applications of computers dates back to the late 1950s, several years before the term "data base" was used in this context. To understand the initial concept of the data base, and its appeal, we must therefore begin by examining the concept of the Management Information System (MIS).

In March 1960, a senior representative of Arthur D. Little, then the largest and longest established management consulting firm, addressed his colleagues at a conference organized by the American Management Association to discuss new applications of computer technology to the problems of corporate administration [1]. Milton D. Stone was, like many of his fellow speakers, enthusing about the incredible potential of the Management Information System, then a very new and very exciting concept [2]. MIS, a concept unveiled to the managerial public for the first time only a year later, was already well on the way to becoming the single most widely discussed concept in the corporate computing world of the 1960s – promoted relentlessly by consultants, "systems men" (corporate staff specialists in administrative management), computer experts and computer manufacturers. Its advocates suggested that the best use of the computer, the only one to truly exploit its potential, was to build an enormous automated system capable of providing to each and every manager in an entire corporation every last piece of information necessary for the performance of their duties, in a timely fashion. It would reach, as Stone put it, "from board chairman to straw boss", and include sophisticated modelling and forecasting capabilities as well as simple factual reporting [1, page 17].

It was men such as Stone who first introduced managers to the idea of information as a generalized, abstract entity, separate from the forms, reports, files and memos in which it had previously been embodied. Stone recognized that a flexible and complete MIS could only be constructed if a firm's entire mass of paperwork could be computerized and integrated "to produce an interrelated body of useful data, or information." He suggested that "this body of data, a veritable 'bucket of facts,' [was] the source into which information seeking ladles of various sizes and shapes are thrust in different locations" [1, page 17]. Others, working with similar ideas, came up with other phrases over the next few years. Another consultant suggested that the office of the future would revolve around a "data hub", defined as "a central source of

information that can serve as an instant inquiry station for executives who need data for decisions.” [3] Representatives of Shell Oil spoke of the need for an “electronic data bank, or pool of information, from which reports of many types can be drawn.” [4, 5].

By the late 1960s, however, “data base” was a common expression in corporate computing circles, largely replacing the hubs, buckets and pools in which data had previously been rhetorically housed. This term was imported from the world of military command and control systems. It originated in or before 1960, probably as part of the famous SAGE anti-aircraft command and control network. SAGE [7] [6] was far more complex than any other computer project of the 1950s, and was the first major system to run in “real-time” – responding immediately to requests from its users and to reports from its sensors. As a result, SAGE had to present an up-to-date and consistent representation of the various bombers, fighters and bases to all its users. The System Development Corporation [8], a RAND Corporation group spun-off to develop the software for SAGE, had adopted the term “data base” to describe the shared collection of data on which all these views were based.

SDC actively promoted the data base concept for military and business use. Its interest in general purpose data base systems was part of its attempt to find new markets for its unique expertise in the creation of large, interactive systems. During the late 1950s and early 1960s, SDC held by far the world’s largest concentration of programmers with experience in large-scale, real-time systems [9]. It paid particular attention to the fashionable area of “time-sharing” computer systems, in which one computer was used interactively by several people, each free to run whatever programs they required. Because computers were then large and expensive, time-sharing promised to make general-purpose, interactive computer use by non-specialists a commercial reality for the first time. SDC invested heavily in this area [10], and identified “computer-centered data base systems” as a key application of time-shared systems – hosting (in collaboration with military agencies) two symposia on the topic in 1964 and 1965. [11].

The approximately 185 participants at the second symposium included high-ranking military officials, business data processing celebrities, and corporate and academic researchers. Reporting on the event in *Datamation*, the leading trade magazine of business computing, Robert V. Head observed that data bases had already unleashed the “biggest single strike” of new jargon “since the great time-sharing goldrush of 1963, ” leaving potential users “sullen and down-trodden.” He concluded by wondering whether it was “possible that users, led by the military, will surrender to these data base systems without a shot being fired in anger.” [12, page 41]

It was around this time that the “data base” term made its first appearances in the ongoing discussion of management information systems. In 1965, Harvard accounting professor John Dearden was using the term “data base” to describe the truly important set of corporate facts and figures that had to be shared between different areas within a business [13]. Within the more technical literature it appeared as a means of pooling information from different files, so that each piece of data would be stored only once. Its great advantage would be “to permit categories of information to be added,

deleted, expanded and otherwise revised, without completely redesigning the file or reprogramming the retrieval routines" [14, page 4].

The idea of the data base as a physical pool of data underlying an MIS was given an early, clear and highly influential statement by Head, who defined the data base as the bottom level of a pyramidal structure [12]. The data base pooled information from all the company's operational systems, and on top of it were erected reporting systems and models to inform higher level managers. [2, 45-50]. The metaphor fit very nicely with the idea of a data *base* supporting the rest of the information system. This obviated the need for systems experts to determine in advance exactly what information each manager would require. Instead managers could interrogate the data base and receive whatever information they needed. The data base was often called a "reservoir" of information [15, 16, 17, page 30].

3. FILE MANAGEMENT SYSTEMS AND DATA PROCESSING

Besides the rather ill-defined concept of the "data base" the other main intellectual ingredient of the Data Base Management System, and the key technological foundation for the actual data base management systems of the 1970s, was the "file management system" (together with its close relation, the "report generator"). File management systems were intended to reduce the cost of producing routine administrative programs, and to make the finished programs easier to change and maintain. Report generation systems made it easier to produce printed reports based on particular criteria.

4. RANDOM ACCESS STORAGE

These file management techniques were very useful with tape storage, but when firms began to start storing their data on disk drives, the extra complexity of programming random access data storage and retrieval made their use almost essential. The disk drive was first offered as a standard option for most major computer systems in 1962 [2] though it had been available in a handful of IBM systems a little earlier. Whereas tape had previously been the only way of magnetically storing reasonably large files of information, it was suddenly possible to hold up to one billion characters of data on the disk drives connected to a single large IBM computer.

5. THE DATA BASE MANAGEMENT SYSTEM AND THE DBTG

The technological innovation represented by systems such as IDS was paralleled by conceptual developments. Until about 1968, the concepts of data bases and file management systems remained largely distinct. The data base was used interactively on-line, could be used by non-specialists and was closely associated with the MIS and the idea of a single huge reservoir of corporate information. File management systems were used primarily by programmers, to reduce development and maintenance costs for routine data processing applications. The most advanced file management systems were beginning to add features to make it easier to pool information from multiple

files, and efforts were underway to add on-line access [17].

As its name suggests, the DBMS was intended to be a new kind of product, extending the capabilities of existing file management systems to support the kind of advanced, on-line, interactive capabilities and huge integrated data stores associated with the data base concept. This was, in many ways, the endpoint of a natural evolution. The DBTG was dominated by the same manufacturers who were adding features to their file management systems and had begun to promote them as supporting, or even being, Management Information Systems [10]. The purpose of the DBTG was to define the capabilities of these new systems, and to develop new standards for them. Its creation was prompted by the realization within CODASYL that COBOL, while doing a great deal to standardize data storage on tape systems and to separate record definitions from program logic, was entirely inadequate when faced with the challenge of random access, disk based storage [10]. On its formation in October 1965 the DBTG had originally been called the List Processing Task Force (its name was changed only in 1967).

6. EARLY DBMS SYSTEMS IN USE

The DBMS enjoyed considerable practical success during the 1970s. By the end of the decade, most large computer installations had installed a DBMS package of some kind. Many of the most financially successful products of the independent software industry were DBMS or file management packages. Adoption of data base management software proved to be a boon to application programmers. In administrative applications of the kind traditionally carried out by corporate data processing departments, an enormous amount of programmer time was taken up doing the things that DBMSs were supposed to automate. They made programs cheaper to develop, much easier to maintain, and facilitated the integration of different business tasks. Data base management technology as defined by the DBTG was very good at dealing with very uniformly structured, hierarchical data of the kind found on administrative forms.

Yet the DBMS never quite lived up to the expectations of people like Nolan, who saw it as a managerial panacea. Indeed, the managerial hype that developed around DBMS technology may have made it hard for firms to make informed technical decisions. As early as 1973, a report [15] by two Booz, Allen & Hamilton consultants suggested that both software and the hardware needed remained immature, that little experience so far existed in its use and that the generalized features offered by the DBMS brought a hefty performance penalty and might well trigger the purchase of more memory or a new processor unit. Most of the true costs were hidden, particularly the staff requirements. As they put it, "Some DBMSs are as complex as the operating system which services them. Also, this group must continuously apply and test new program fixes and new features to keep the system 'alive and well.' It is not uncommon to see a small systems programming team double or even triple as the result of a DBMS" [13, page 74]. Later reports, in [17], suggest that these problems continued for several years, and that many firms installed DBMS packages because of a "bandwagon" rather than a careful and informed evaluation.

7. THE DATA BASE MANAGEMENT SYSTEM SINCE 1980

In 1973, Charles W Bachman was awarded the Association for Computing Machinery's Turing Medal – the most prestigious award in computer science. The citation singled out his creation of the pioneering IDS system (which it retroactively termed a DBMS) and his work on the DBTG to incorporate these ideas into its specifications. This award was in itself an important event, representing a new level of acceptance among computer science researchers of data base problems as intellectually respectable subjects of inquiry alongside better established areas such as numerical analysis, compiler theory and the theory of algorithms. The event is better remembered, however, for Bachman's speech [7]. Entitled "The Programmer As Navigator, " it developed the idea that the shift to DBMS technology represented something akin to the Copernican revolution – in that the work of programmers would now revolve around the data base rather than the hardware of the computer. Though this prophecy took several decades to come true, knowledge of data base systems has now become a fundamental requirement for virtually all administrative applications programming, systems analysis and advanced web design work. But, as its title also implied, the impact of generalized DBMS would be much greater for programmers than for managers.

The acceptance of the DBTG concept of a data base management system thus implied a new and more concrete vision of what a data base was – basically a body of electronic data that could be managed by a data base management system. As such, the commercial success of DBMS packages supported the growing prestige of corporate computing staff, against attempts by information scientists and documentationalists [5] to turn the library, rather than the computer room, into the heart of any corporate information system. Despite the MIS influenced hopes of the 1970s that a DBMS could be the heart of a system including all corporate information, it proved adept at handling only a small subset of this material.

11. REFERENCES

- [1] Stone, M. D. Data Processing and the Management Information System: A Realistic Evaluation of Data Processing's Role in the Modern Business Enterprise. In American Management Association ed. *Data Processing Today: A Progress Report--New Concepts, Techniques and Applications--AMA Management Report Number 46*. American Management Association, Finance Division, New York, 1960, 14-22.
- [2] Haigh, T. Inventing Information Systems: The Systems Men and the Computer, 1950-1968. *Business History Review*, 75, 1 (Spring 2001), 15-61.
- [3] Weindling, R. E. Office Will Run Every Business Activity. *Office Management and American Business*, 22, 1 (January 1961), 12-15.
- [4] Haslett, J. W. Total Systems-A Concept of Procedural Relationships in Information Processing. In Meacham, A. D. and Thompson, V. B. eds. *Total Systems*. American Data Processing, Inc. , Detroit, MI, 1962, 16-19.
- [5] Keller, A. E. The Man Behind Systems at Shell Oil. *Business Automation*, 7, 2

- (February 1962), 20-24.
- [6] Edwards, P. *The Closed World: Computers and the Politics of Discourse in Cold War America*. MIT Press, Cambridge, MA, 1996.
 - [7] Hughes, T. P. *Rescuing Prometheus*. Pantheon Books, New York, 1998.
 - [8] Baum, C. *The System Builders: The Story of SDC*. System Development Corporation, Santa Monica, 1981.
 - [9] Rowan, T. C. The Recruiting and Training of Programmers. *Datamation*, 4, 3 (May-June 1958), 16-18.
 - [10] Anonymous A panel Discussion on Time-Sharing. *Datamation*, 10, 11 (November 1964), 38-44.
 - [11] System Development Corporation *Preprint for Second Symposium on Computer-Centered Data Base Systems*, Sponsored by SDC, ARPA, and ESD, September 1 1965, Burroughs Corporation Records (CBI 90), Charles Babbage Institute, University of Minnesota, Minneapolis.
 - [12] Head, R. V. Data Base Symposium. *Datamation*, 11, 11 (November 1965), 41.
 - [13] Dearden, J. How to Organize Information Systems. *Harvard Business Review*, 43, 2 (March-April 1965), 65-73.
 - [14] Simon, L. and Sisson, R. Evolution of a Total System. *Total Systems Letter*, 1, 11 (January 1966), 1-4.
 - [15] Head, R. V. MIS-II: Structuring the Data Base. *Journal of Systems Management*, 21, 9 (September 1970), 37-38.
 - [16] Kircher, P. Breakthrough in Management Information Systems. *Journal of Data Management*, 7, 2 (February 1969), 28-31.
 - [17] Wendler, C. C. What are the Earmarks of Effective Total Systems? *Systems & Procedures Journal*, 17, 4 (July-August 1966), 29-31.

