

Abstract for History and Heritage of Science Information Systems

“A Veritable Bucket of Facts”: Origins of the Database Management System

The Data Base Management System (DBMS) provides a vital underpinning for today’s scientific and corporate information systems. No digital library, scientific database or search engine could function without database technology. Yet commercial DBMS have been geared to the highly structured information needs of corporate administration, rather than the full-text searching and content-centric needs of many scientific information systems. As sold during the 1970s they required expert programming, were batch-oriented and inflexible. Even relational systems, the dominant paradigm taught in current information systems courses, are far better suited to record keeping than to research. While many pioneering database systems tackled scientific problems, as fundamental database technology was packaged and standardized via the DBMS most large scientific applications continued to rely on expensive custom development. Only the recent rise of the web has brought the challenges of massive, loosely structured and distributed databases closer to the commercial mainstream.

Why did DBMS technology take this path? I explore this question by examining the origins of the DBMS concept and of the database concept itself. Surprisingly, the two turn out to have quite different origins. The database concept (originally “data base”) derives from early military on-line systems, such as those constructed by the System Development Corporation (SDC). I use archival research to explore SDC’s work in the area, its extension of the concept to scientific systems and its attempt to commercialize database technology as part of timesharing services. I also document the role of Informatics, another firm heavily active in the development of on-line information systems under government contract. While the idea of an integrated database spread into corporate data processing and “systems” circles during the early 1960s it was seldom, if ever, realized in practice.

The DBMS, however, was essentially a re-labeling of an earlier technology: the generalized file processing system. Generalized file and list management routines were among the first utility programs to be produced by firms like General Electric as they put administrative computers to work during the mid-1950s. These routines were some of the first systems tools to be packaged and distributed by computer vendors, and were also perhaps the most important product of the nascent packaged software industry of the 1960s. Only during the very late 1960s were these products re-christened as “Data Base Management Systems” (indeed, the Data Base Task Group, the industry group that popularized the term, began its life as the List Processing Task Group). While considerable attention was given to the ideas of on-line access and end-user querying, in practice leading DBMS systems of the 1970s, such as Informatics Mark IV, were used almost exclusively by programmers and in batch mode.

As the DBMS concept spread, the older idea of the database itself was largely redefined to fit this new model – as the thing managed by a DBMS. Its adoption reflected and reinforced the gulf between scientific information problems in indexing, information retrieval, and taxonomy and standard commercial tools for data management. Within the corporation, responsibility for database matters fell to the corporate computing or MIS staff, rather than librarians or information scientists. As well as documenting the origins of the DBMS, itself a vital and little-explored topic, my paper also serves to illuminate these broader issues in the social and organizational history of information.