

“A Veritable Bucket of Facts”

Origins of the
Data Base Management System 1960:1975



Tom Haigh – thaigh@sas.upenn.edu

My Topic:

- Origins of the database management system (DBMS)
 - Most important class of corporate IT infrastructure
 - Foundation of web, e-business
- Part of broader project on corporate computing
 - Focus on use of technology
 - Professional, organizational, managerial issues

Structure of Paper

- Skeleton of written version
 - Draft available from www.tomandmaria.com/tom
- Four sections
 1. Origins of Data Base concept
 - Cold war military, Information science related
 2. Origins of file management system
 - Corporate data processing – clerical routine
 3. Early discussion of data bases for business
 4. The DBMS
 - The data base meets the file management system

The Data Base Concept

Section 1

The Term Data Base

- Data base concept of military system origin
 - Probable source is System Development Corporation (SDC), 1960 or earlier
 - Predates the DBMS by almost a decade
 - SDC had software contract for SAGE project

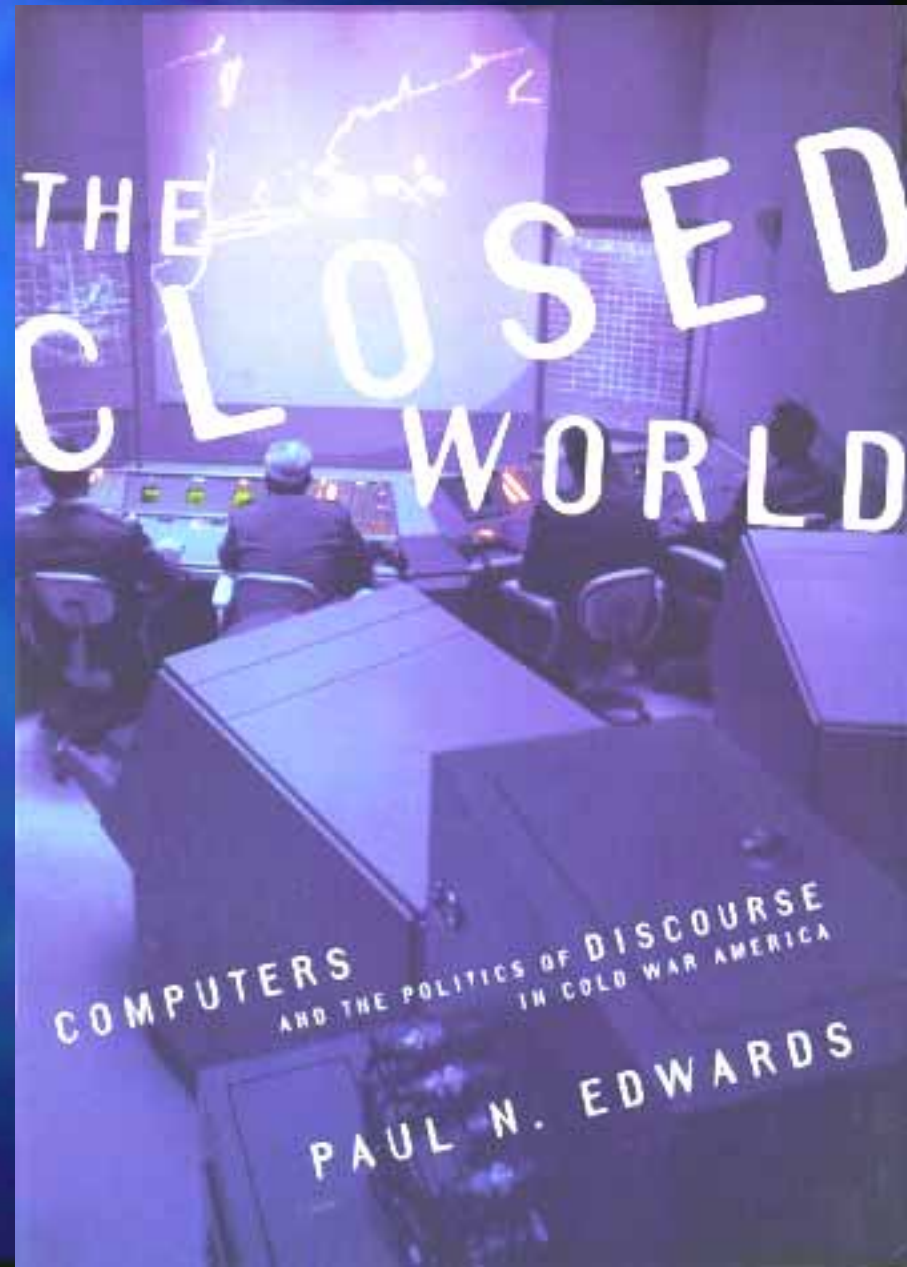
A "Semi-Automated Ground Environment"!

- SAGE itself was an anti-bomber air defense network in 1950s & 1960s
- Highly automated system
 - Collects data from huge network at central command posts
 - Decisions made very rapidly
- Enormously expensive
 - Most important single project in history of computing

The Closed World

- Cultural history of the SAGE air defense system and the SDI project

Edwards, Paul. *The Closed World: Computers and the Politics of Discourse in Cold War America*. Cambridge, MA: MIT Press, 1996.



Data Base in SAGE

- Shared repository of data
- Crucial Characteristics
 - Constantly updated
 - Accessed interactively (“real-time”)
 - Data base is shared between users/systems, gives different views to each
- SDC develops interest in “Information Retrieval”

Information Retrieval

- New concept circa 1950
 - New technologies & techniques for searching data
 - Tied to cold war “information explosion”
 - Increasingly associated with computer & electronics
- Contemporaneous with
 - Information Theory (late 1940s)
 - Information Science (coined 1959?)
 - Information Technology (1958)
- Discussion of information in generalized way is new, particularly to business

SDC Tries to Commercialize

- Early-mid 1960s:
 - funding work in information retrieval
 - unique expertise in on-line systems, time-sharing
- Pioneer “computer centered data base systems” for administrative uses
 - LUCID (on-line “data management system” for non-programmers)
 - Finds some governmental use, leads to TDMS
- Late 1960s
 - Timesharing/computer utility concept
 - 1968: SDC launches CDMS nationally. Huge flop

On-Line IR in the 1970s

- Market for on-line Information Retrieval grows in bibliographic niches in 70s
 - SDC turns air-force systems into ORBIT
 - Lockheed builds RECON document management system for NASA, basis for later DIALOG commercial service
 - Informatics turns reworked RECON into POPINFO, TOXLINE, ENVIRON for Feds.
- RECON IV flops as commercial package
 - Public sector service; not private product

The File Management System

Section 2

The Electronic Era for Business



Data Processing Tasks

- Payroll, accounting, invoicing
 - Taking over jobs from existing punched card machines
 - Slow evolution hardware of hardware, practice
- Intended to automate clerical work
 - Success means replacing clerks
 - Justified on basis of lower operating costs

File Management Software

- As old as corporate computing
 - First documented in GE, mid-1950s
 - Generalized set of subroutines to update, query, maintain sequential files
- By mid-1960s, becoming more sophisticated
 - Offered as commercial products
 - Working with new random-access devices
- Mark IV (Informatics) is huge success
 - Also IDS (GE), IMS (IBM)

Data Base enters managerial discussion

Section 3

"Data base" in corporate use

- "Data base" concept crosses over to corporate use in early 1960s
- "Total" Management Information System
- Hugely popular idea in 1960s
 - Integrated reporting and control systems
 - All data for all managers
 - Interactive use in real-time
 - Spans entire firm
- Impossible to achieve

Data Base: Early Mgmt Usage

MIS relies on a

“body of data, a veritable ‘bucket of facts,’ [as] the source into which information seeking ladles of various sizes and shapes are thrust in different locations.”

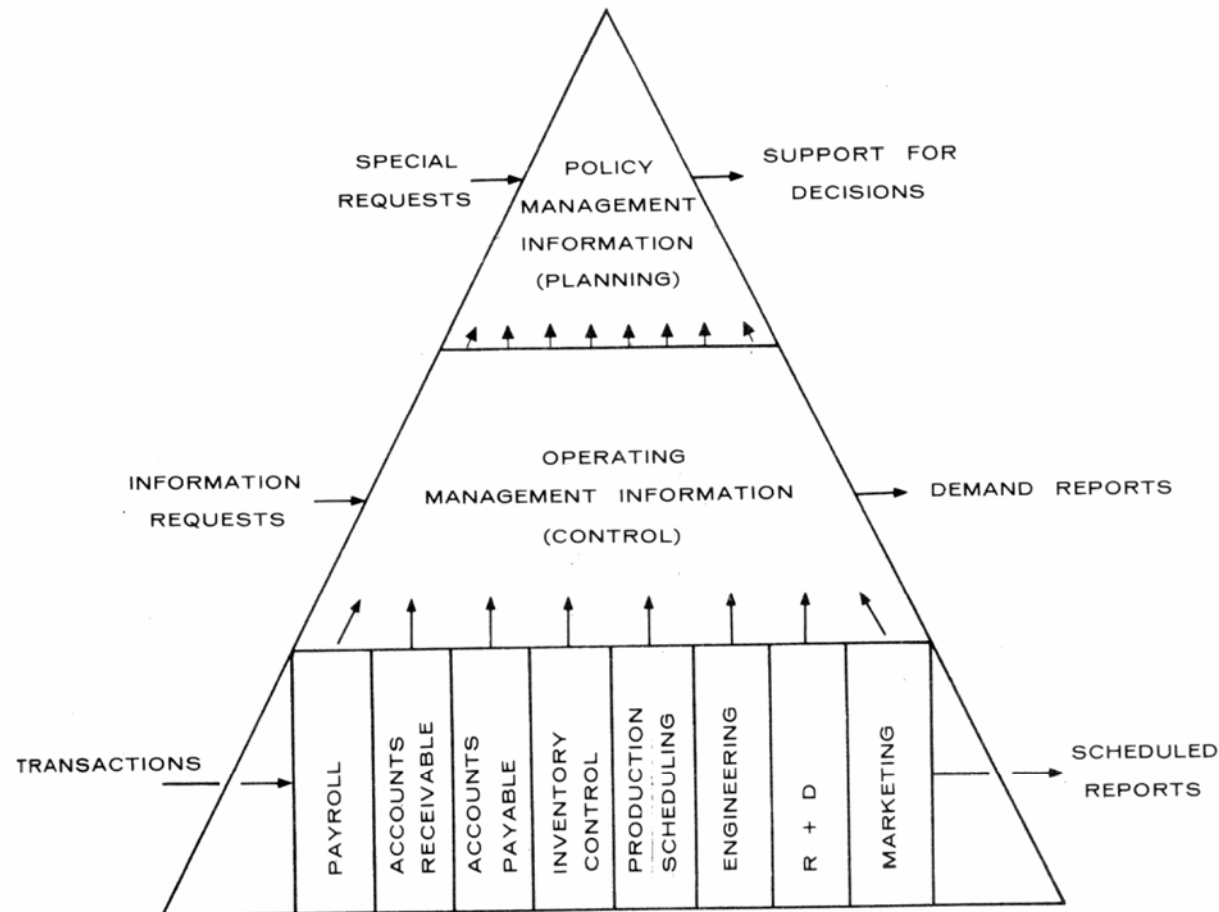
(Milt Stone, 1959)

- Variations in 1961/1962:
“data hub”, “data bank”, “pool of information”
- “Data Base” spreads in mid-1960s

The Information Pyramid (1967)

- “Information” ties together all levels of management & operations
- Bottom level of the pyramid is the “data base”

Fig. 1. Management Information System



State of Play circa 1967

- Data base concept is
 - Fashionable
 - Widely promoted as key to MIS
 - Vaporware, revolutionary
 - Real-time, on-line, "total system"
 - Closely tied to information retrieval
- File management software is
 - Growth area
 - Data processing tool (batch mode)
 - Practical, batch-oriented, evolutionary

The Data Base Management System

Section 4

The DBMS & CODASYL

- New concept “Data Base Management System” appears circa 1968
 - CODASYL Data Base Task Group
 - Originally in context of extensions to COBOL
 - Based on consideration of current file management products, directions for future.
- One system must offer
 - Real Time & Batch operation
 - Capabilities for programmers
 - Ability to query directly

DBMS – Foundational Concept

- DBMS as software layer between data, users
 - Different interfaces, languages for
 - Programs & programmers
 - Ad-hoc managerial reporting
 - Data definition
 - maintenance and administration
- Sets up links between files
 - BUT rigid, standardized format remain

DBMS as a Product

- Term DBMS applied widely to new & existing products
 - CODASYL standard influential but not dominant
 - Guides evolution of packages
- DBMS key part of software industry
 - TOTAL, IDMS, SYSTEM 2000, IMS (IBM)
- Even in late 1970s, used mostly in batch mode
 - Real-time very inefficient
- Big cost in hardware and software
 - New specialists needed to configure

DBMS usages in the 1970s

- Advantages mostly for programmers
 - easier reporting,
 - Program/data independence
 - faster application development,
 - easier maintenance
 - better integration of different applications
- Integration proves harder than expected
- Help with conversion to disk and multitasking operating system

Hopes for MIS reborn with DB

- "Writings on MIS have waned recently and have largely been replaced by writings on the Data Base" (1973)
- The "Data Base Administrator"
 - Originally expected to take responsibility for "data as a resource... much broader than machine readable data" (1974)
 - "something of a superstar" (1975)
- DBMS technology expected to build integrated, company wide DB

Post 1980: DBMS Concept Spreads

- Shift to relational model
 - Devised in 1970s, spreads in 1980s
 - SQL emerges as standard
- Costs lower, performance improves
 - But still tool mostly of new programmers
- Extension to new kinds of hardware
 - Minicomputers
 - Microcomputers
 - Pocket computers!

DBMS as Information Technology

- Compared to 1960s data base ideas
 - New concept of database is narrower
 - More general information retrieval problems are excluded
- DBMS is not well suited for
 - Irregular records
 - Full text or even keyword searching
 - Ad-hoc linkages between records
 - Context, relevance (in IS terms)
- Only with search engines of 90s
 - Is much attention given to unstructured data

Implications

- Despite IR, IT, etc. hard to deal with information in general
 - Routine administrative (dominant in business use) – file management, DBMS
 - Scientific and bibliographical (library) – specialized on-line system
- In practice, data bases fragment
 - New challenge is reuniting them!
- New dreams of integrated systems
 - Data warehouse (reporting)
 - Enterprise Resource Planning (operational)₂₉

More on my Website

- www.tomandmaria.com/tom
- Papers, including
 - Full draft of this one
 - "Inventing Information Systems" on MIS
 - "The Chromium-Plated Tabulator" on data processing
- Computer history resource guide