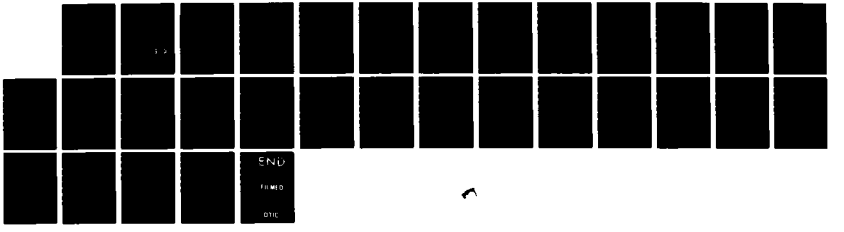
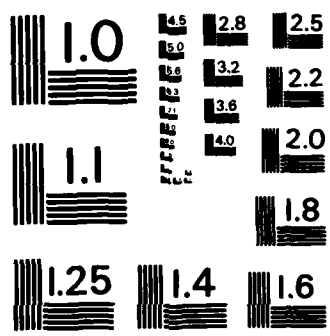
AD-A162 203 DEVELOPMENT OF A DATABASE DESIGN FOR SERIALS CONTROL IN 1/1
THE DEFENSE COMMUNICATIONS AGENCY(U) DEFENSE
COMMUNICATIONS AGENCY WASHINGTON DC TECHNICAL AND MAN..
UNCLASSIFIED D A GUERRIERO 21 OCT 85 F/G 5/2 NL





(2)

AD-A162 203

DEVELOPMENT OF A DATABASE DESIGN FOR SERIALS CONTROL IN THE DEFENSE COMMUNICATIONS AGENCY

by

Donald A. Guerriero

Defense Communications Agency
Personnel and Administration Directorate
Employee Development Division
Technical and Management Information Center

Washington, DC 20305

OCTOBER 21, 1985

TABLE OF CONTENTS

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

AD-A162 203

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT UNLIMITED
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Defense Communications Agency 6b. OFFICE SYMBOL (If applicable) H396	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (City, State and ZIP Code) Defense Communications Agency Technical and Management Information Center Washington, D.C. 20305	7b. ADDRESS (City, State and ZIP Code)
8a. NAME OF FUNDING/SPONSORING ORGANIZATION Defense Communications Agency 8b. OFFICE SYMBOL (If applicable) H396	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (City, State and ZIP Code) Defense Communications Agency Technical and Management Information Center Washington, D.C. 20305	10. SOURCE OF FUNDING NOS. PROGRAM ELEMENT NO. / PROJECT NO. / TASK NO. / WORK UNIT NO.
11. TITLE (Include Security Classification) SEE REVERSE	
12. PERSONAL AUTHOR(S) GUERRIERO, Donald A.	
13a. TYPE OF REPORT Final 13b. TIME COVERED FROM 850601 TO 851021	14. DATE OF REPORT (Yr., Mo., Day) 851021 15. PAGE COUNT 28
16. SUPPLEMENTARY NOTATION	

17. COSATI CODES		
FIELD	GROUP	SUB. GR.
5	2	

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
SEE REVERSE

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

SEE REVERSE

20 DISTRIBUTION/AVAILABILITY OF ABSTRACT UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS ☐	21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL GUERRIERO, Donald A.	22b. TELEPHONE NUMBER (Include Area Code) (703) 692-0373 22c. OFFICE SYMBOL H396

DD FORM 1473, 83 APR EDITION OF 1 JAN 73 IS OBSOLETE.

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

11: TITLE: Development of a Database Design for Serials Control in the Defense Communications Agency

18: SUBJECT TERMS: Information Science, Information Retrieval, Libraries, Library Automation, Serials, Serials Control, Databases, Periodicals

19: ABSTRACT:

The Defense Communications Agency subscribed to about 500 serials, including technical journals, magazines newspapers, government publications, microform editions of technical and administrative publications, and looseleaf services. Within the Agency the Technical and Management Information Center was responsible for controlling all subscriptions for the Agency. Using existing manual systems did not provide effective control over the subscriptions. Studies had shown that an automated system was both possible and desirable from a technical and management point of view. This project was designed to develop an inhouse database for serials control in the Defense Communications Agency. The purpose of the project was to develop the logical structure for the database and identify the software and hardware to use for this database. The procedures were simply to gather the available information for the database, determine the objectives for the database, describe the organization of the data in a formal way, and to determine what software and hardware could be used to enter, store, and retrieve the information. This study showed the logical structure of the proposed database and the requirements for implementation. Recommendations were made to implement the database immediately using an IBM PC (M300 version) and dBase III. Benefits and costs of the database were studied and it was predicted that the database should be cost-effective.

ABSTRACT

The Defense Communications Agency subscribed to about 500 serials, including technical journals, magazines newspapers, government publications, microform editions of technical and administrative publications, and looseleaf services. Within the Agency the Technical and Management Information Center was responsible for controlling all subscriptions for the Agency. Using existing manual systems did not provide effective control over the subscriptions. Studies had shown that an automated system was both possible and desirable from a technical and management point of view. This project was designed to develop an inhouse database for serials control in the Defense Communications Agency. The purpose of the project was to develop the logical structure for the database and identify the software and hardware to use for this database. The procedures were simply to gather the available information for the database, determine the objectives for the database, describe the organization of the data in a formal way, and to determine what software and hardware could be used to enter, store, and retrieve the information. This study showed the logical structure of the proposed database and the requirements for implementation. Recommendations were made to implement the database immediately using an IBM PC (M300 version) and dBase III. Benefits and costs of the database were studied and it was predicted that the database should be cost-effective.

Introduction

The Defense Communications Agency (DCA) performs systems engineering for the Defense Communications System (DCS) to insure that the DCS is planned, improved, operated, maintained, and managed effectively, efficiently, and economically to meet the long-haul, point- to-point, and switched network telecommunications requirements of the National Command Authorities, the Department of Defense (DoD), and other Federal agencies. Secondly, DCA provides system engineering and technical support to the National Military Command System (NMCS), Minimum Essential Emergency Communications Network (MEECN), and the Worldwide Information Network (WIN). Finally, the Agency performs system architecture functions for Military Satellite Communications and ADP support to the joint Chiefs of Staff, Secretary of Defense, and other DoD components.

Within DCA the Technical and Management Information Center supports more than 3500 Washington, D.C. area DCA employees with information from a variety of sources, including periodicals and other serial publications. The Agency currently subscribes to about 500 government and commercial serials, in hardcopy and microfilm, through a variety of acquisition channels. The

estimated total annual cost for the subscriptions is about $130,000. According to current DCA regulations, The Technical and Management Information Center is responsible for ordering subscriptions for the Agency. A recent DCA Management Engineering Team (MET) study concluded that there was inadequate control over the number and type of subscriptions purchased by DCA. The Technical and Management Information Center was given the requirement to study ways and means to control subscriptions.

Background and Significance

The Technical and Management Information Center management was aware of this problem for some time. During the past two years several actions were taken to obtain more accurate data on the subscriptions being ordered and to either cut down the cost of the subscriptions or at least assure Management that all those being ordered were needed. In addressing this problem, two studies were conducted in 1983, 1984 to (1) determine the feasibility of using a computerized serials control system in DCA and (2) to study the relevance of the certain periodicals now purchased by DCA. These studies and the intuitive feelings of the Center's staff and Director showed that a computerized system for serials control is feasible and necessary.

As a result of the first study, three possible methods of automation were considered. The vendor system, EBSCONET, was contracted for under a pilot program. EBSCONET is now being used to provide the Center with a listing of periodicals ordered through EBSCO, the Center's serials agent, but the system does not provide all the data on serials needed by the staff. The OCLC Serials Control subsystem was to be used as well in the pilot study until OCLC changed the serials control system to one that requires a substantial initial investment in equipment. The Center was unable to secure funds for this equipment and did not

use the OCLC system. A third alternative, creation of an in-house serials control system was also attempted, but failed when it became evident that the existing word processing equipment in the Center was inadequate for storing and manipulating extensive records of subscriptions. This project was begun to design an in-house database that would use the existing computer facilities of the Agency and that would include all the information needed by the staff on the subscriptions.

The Technical and Management Information Center management continued to feel the need for an automated system that would provide adequate information for serials management. The purpose of this project was to develop a written report on the identification of available information for a serials database and creation of a data entry form. The logical structure of the database was to be determine after studying the objectives of the database creation.

Additionally, hardware and software available in DCA for database management was to be identified. A schedule for design of the database, testing and evaluation, implementation, and data entry was to be developed in consulation with the DCA Office Automation group.

According to Blair (1982:15) the most critical part of creating a database is the determination of exactly what is going to go into a record and in what order the data will be entered. This study concentrated on the former in the belief that, the data in the database would determine its usefulness and success. Again, following Blair's advice, there was consultation with all those who would use the data on what fields would satisfy their needs. Following Martin (1977:23) an effort was made to develop a database that would be able to change and grow.

The complexity of serials control, little understood outside the library, needed to be taken seriously. Kuhns (1976-1977:173:181) pointed out the difficulties inherent in serials librarianship. The expectation was that changes in use and uncertainties about software and hardware would require flexibility. Careful attention was given to the primary and secondary objectives of database organization listed by Martin (1977:44-45). Although the Technical and Management Information Center uses the services of Agency and DoD supply services and subscription agencies, the need for internal control at the operating level was critical. While Huff (1976:690) pointed out that serials agents have the primary goal of service through ordering, claiming, and renewing, the Center needed better information to secure this service. The arrival of the microcomputer and database management systems for this machine

provided new opportunities for developing databases with or without the services of a centralized systems departments. Riley (1981:153) wrote on the poor record of systems developers in meeting schedules, conforming to budgets, and producing satisfied users. The need for adequate planning was shown, but today the record of systems departments is no better. Lacking this support the management of the Technical and Management Information Center determined to try to develop an inhouse system. Adequate information on database design was available in many texts and articles including Hussain & Hussain (1981: 118-213), Martin (1977:22-291) and Ullman (1983:1:34 and 145-172) among many others. User-oriented periodicals also have highlighted the availability of database management systems on microcomputers as in Krajewski (1984:135-212). The availability of low cost storage and database management systems for microcomputers provide new possibilities for information users.

Procedures

The first requirement in this project was to determine the information required and available for the database. Using existing manual and computerized records a consolidated list of all periodicals ordered by DCA Information Centers was first developed. However, this list included only title and holdings information. Additionally, the serials agency, EBSCO, provided lists of all periodicals ordered by DCA. This was important in that the DCA Information Centers are not the sole ordering entity within the organization. Again, however, these lists provided only titles, and some additional data pertinent to EBSCO orders.

Important criteria for the database were determined to be the following:

(1) Designed for implementation on the hardware and software now available in DCA with the possibility of migration of the database to other hardware and software as it becomes available.

(2) If possible, designed for multiple access by users at different DCA locations with adequate controls on access.

(3) Include only that data necessary to provide management information required by DCA and the Technical and Management Information Center.

(4) Format similar to MARC cataloging format for serials.

(5) Format to provide ease of data entry and change.

(6) Expansion of the database to cover serials, other than periodicals, should be possible.

(7) Procedures for information retrieval and report generation should be simple.

All the above criteria and the availability of information were considered. With that information in hand a format for the database was developed.

The question of (1) designing for implementation on the hardware and software available now in DCA with the possibility of migration of the database to other hardware and software as it becomes available and (2) multiple access by users at different DCA locations, with adequate controls on access, was discussed with the staff of the DCA Office Automation group.

Results

A logical organization of the data was developed using tabular representation. According to Martin (1977: 202-229), data can be described so that it can be represented to nonprogramming users in a way that it can be easily understood. Additionally, new data items, records and associations can be added to the database without changing the existing subschemas or changing application programs. Maximum flexibility in handling data can be obtained. The users view of the data can be kept separate from the physical representation of data and the hardware can be changed.

This was important in this project, due to the uncertainty of the eventual software and hardware to be used for the database. The method of representing data for this project was to use the two-dimensional format of a relational databases. This required normalization of the database and the construction of a relational database. Data representation in the relational database is as tables showing "relations" and "attributes." Another advantage of this procedure is to represent the data, to users, not trained or talented in the techniques of data processing, in an understandable way. By so doing adequate feedback can be obtained from users on the applicability of the data being considered for use.

As a result of investigation of the data requirements the following tables showing relations and attributes were developed. Each relation is shown with its relation name and its attributes. Each attribute is shown with its name, type, length, and description.

LOGICAL DESCRIPTION OF THE PROPOSED DATABASE

Relation: Description

Attribute Name	Type	Length	Description
title	c	60	The title of the serial
title2	c	60	For the rare title that might have more than 60 characters
title3	c	60	Alternate title
type	c	2	A one-character code identifying the type of record. This is determined by the type of item being entered and could be either P for printed material or M for an original microform publication.
confpub	c	2	A one-character code that shows whether the work contains proceedings, reports, or summaries of a conference, meeting or symposium. 0 (zero) = not a conference publication 1 = a conference publication

sertyp	c	2	A-one character code that shows whether the serial is a periodical, monographic series, or newspaper. p = periodical m = monographic n = newspaper 0 (zero) = none of the above
phymed	c	2	A one-character code that specifies the physical form in which the serial is originally published as opposed to that which it is when it is entered in the database. 0 (zero) = None of the following a = microfilm b = microfiche e = newspaper format z = other

cont	c	2	A one-character code telling what a serial contains (as opposed to what it is). One, two, or three codes may be included. If fewer than three codes are used, the values should be left-justified (followed by blanks). If more than three codes are applicable, use the sequence of the list below to determine which three to include. a = abstracts i = indexes o = book reviews b = bibliographies c = catalogs r = directories y = yearbooks s = statistics l = legislation w = law reports and digests g = legal articles v = legal cases and case notes h = biographies d = dictionaries e = encyclopedias p = programmed texts f = handbooks 0 (zero) = no specified contents

freq	c	2A	one-character code covering the frequency of publication d = daily i = three times a week c = semiweekly (twice a week) w = weekly j = three times a month e = biweekly (every two weeks) s = semimonthly (twice a month) m = monthly b = bimonthly q = quarterly t = three times a year f = semiannual (twice a year) a = annual g = biennial (every two years) h = triennial (every three years z = other u = unknown (used when the frequency is not known and it also is not known if it is "irregular" *0 (zero) = irregular*
pubstat	c	2	A one-character code that shows whether the serial currently being published. c = currently published d = dead u = unknown s = suspended

Relation: number

attribute name	type	length	Description
title	c	60	
oclc	c	20	OCLC Number
coden	c	10	CODEN Number
issn1	c	10	ISSN Number
issn2	c	10	ISSN Number
EBSCO	c	10	Serials Agent (EBSCO) Number
mda	c	20	Invoice Number
account	c	15	Account Number
order	c	10	Order Number

Relation: orderinfo

attribute name	type	length	Description
title	c	60	
vendor	c	10	Code for Vendor

Relation: cost

attribute name	type	length	Description
title	c	60	
cost	i	2	Annual Cost

Relation: codes

attribute name	type	length	Description
title	c	60	
code1	c	10	Internal addresses (all "codes")
code2	c	10	
code3	c	10	
code4	c	10	
code5	c	10	
code6	c	10	
code7	c	10	
code8	c	10	

Relation: holdings

attribute name	type	length	description
title	c	60	
holdings	c	60	Holdings in DCA
locate	c	2	Location in DCA

Relation: checkin

attribute name	type	length	description
title	c	60	
dates	c	120	Dates item received

Relation: claims

attribute name	type	length	description
title	c	60	
history	c	60	Claims made

Relation: renewal

attribute name	type	length	description
title	c	60	
renstat	c	2	(Renewal status (Y or N)
expdate	c	8	(Date of expiration)

Finding software and hardware for the database proved much more elusive. The Technical and Management Information Center had available a Wang System 5 with word processor with sorting, and telecommunication software only. This system did not have a hard disc or other mass storage and could not accommodate the lengthy files necessary for serials control. The Center also had available an IBM PC modified by OCLC as an M300. This machine had dual floppy discs and minimal memory. Again the hardware and software did not satisfy the needs of the serials control database.

The DCA Office Automation Support Center (OASC) suggested four possible alternatives for software and hardware requirements. These were

LIST PROCESSING ON A Wang OIS or Wang VS-100.

List processing is a specially-designed, easy-to-use package that can be used to create and maintain files such as those needed in serials control. The Wang OIS system was available to users in DCA with compatible terminals. However, the Technical and Management Information Center did not have a terminal. Similarly, no terminal was available for the Wang VS-100 that would be put into use in DCA in mid-1986.

Wang Professional Computer with Hard Disc and DBMS.

Four to six weeks acquisition time would be required. Although the microcomputer, hard disc, and DBMS system theoretically could be purchased, there were no unencumbered funds available for this purpose.

Upgrade IBM PC with a hard disc and DBMS

This would require a minimal investment in a hard disc and a DBMS. There was the question of compatibility with the Wang equipment being installed in DCA. The microcomputer-based system would be cumbersome and would not have easy backup. The Technical and Management Information Center staff would be required to develop the database implementation.

Wang VS-100 with ResponseR DBMS

There would be a delay until mid-1986 before the machine would be available and at that time it would already be heavily committed to other applications. There would be the necessity of acquiring a compatible terminal to access the Wang VS-100.

Discussion, Implications and Recommendations

The information for the database can be obtained from readily available sources with minimal difficulty and will meet the immediate needs of DCA and the Technical and Management Information Center. These sources include the following:

The serial in hand
Serials directories
OCLC records
EBSCO catalogs
DCA contracts
DCA financial information
DCA order and renewal records
DCA serials lists

Non-professional staff can be trained to obtain the necessary data and enter it in a simple form devised from the logical description of the database. As a result of the separation of the logical description of the data and the physical representation and the hardware, it will be possible to complete this part of the project before actually entering any data.

The requirements for additional software and hardware is more difficult to solve in the short run. Nevertheless, from the available information it is recommended that a DBMS package for the IBM PC (M300) now in the Technical and Management Center be purchased to begin to implement the database. The package recommended is the dBase III DBMS by Ashton-Tate. This DBMS supports the IBM Personal Computer. It features over a billion records per database; 128 fields per record; variable-length text fields (up to 4K bytes per entry); 4000 bytes per record (fixed length), 512 bytes per record (variable length), 10 database files in use simultaneously; sorting, indexing, and other file operations, and full-screen formatting. dBase III is a relational database management system with interactive data entry. Setting up the database can be done by the current Technical and Management Information Center staff, although it will probably take about 160 hours of professional staff time, an equal amount of non-professional time, and about 48 hours of training on the system for six professionals.

Additional hardware will be required for this system. The hardware recommended and probable cost is:

IBM Personal Computer (M300 version)	$3500
Tallgrass disc drive with tape backup	8000
Maintenance (yearly)	1100
Software	1000
Cables etc	200
Total	$13,800

If sufficient funds are not allocated for the equipment, it is recommended that one 10 MB Winchester hard disc and drive be acquired to use with the existing M300. An additional M300 should be acquired as well. This will lower the cost to about

$6,000 and provide an interim solution. When a permanent solution can be found the equipment will still be useful for many other applications in the Technical and Management Information Center.

Furthermore it is recommended that funds be obtained to develop a multi-user, multi-purpose system for the Technical and Management Information Center. The exact configuration depends on factors that cannot be addressed in this report, but the use of current DCA office automation and computer resources staff as well as outside consultants familiar with specialized systems used in libraries and information centers would be helpful. The estimated cost for a system (in 1986) would probably be from $80,000 to $130,000.

If the recommendations contained in this study are followed it is expected that an operational database for serials control will be available for use by the Technical and Information Center staff by December 30, 1985. The potential for improvement in the operations of the Technical and Management Information Center will include several potential tangible and intangible benefits that should outweigh the costs of the system.

Probably the most important tangible benefit will be increased managerial control over a large expenditure of public funds. The development and implementation of this database is

expected to cost a maximum of about $14,000 for equipment. Staff costs will be about $7,000, and training about $4,000. The total first year costs will therefore be about $25,000. The system will probably be in use for about 5 years, but even if it is not migration of the data to other automated systems will be simplified by this implementation. The equipment has at least a five year life and can be used for other requirements. Training of the staff will also provide benefits since the dBase III software can be use to create other databases.

After the database is available it is estimated that administrative costs of serials control can be reduced by about $2000 per year in labor costs. Better control of the serials will result in either savings in the costs of subscriptions through the elimination of duplicate or unneeded subscriptions. Savings are expected to be about $10,000 in the first year. These savings can be applied to more productive information resources or other uses.

Intangible benefits will include improved customer service as the Technical and Managment Information Center will be able to order, renew, and claim missing issues much faster. Additionally, patrons will be provided with up-to-date information on serials available. Staff morale will be improved through the elimination of cumbersome, tedious, and poorly

designed manual control systems now in use. There will be opportunities to make more informed decisions on subscription retentions and planning for future information requirements. Finally, a critical factor of meeting recurring requirements for information on subscriptions from DoD, GAO, and DCA auditing and inspection teams will be met.

BIBLIOGRAPHY

Blair, John C. "Creating Your Own Database" Database, August 1982, pp. 11-17.

Huff, William H. "Serial Subscription Agencies" Library Trends, April, 1976, pp. 683-694.

Hussain, Donna and Hussain, K.M. Information Processing Systems for Management. Homewood, Illinois: Richard D. Irwin, Inc., 1981.

Krajewski, Rich (ed.) "Databases" Byte, October, 1984 pp. 137-234.

Kuhns, Kathleen. "Serials Librarians and Their Discontents" The Serials Librarian, Winter 1976-77, pp. 173-181.

Martin, James. Computer Data-Base Organization, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, Inc. 1977.

Riley, M.J. Management Information Systems, 2nd ed. San Francisco: Holden-Day, Inc., 1981.

Rowe, Alan J., Mason, Richard O., and Dickel, Karl E. Strategic Management & Business Policy. Reading, MA: Addison-Wesley, 1985.

Ullman, Jeffrey D. Database Systems, 2nd ed. Rockville, MD: Computer Science Press, 1982.

U.S. Defense Communications Agency. Five Year Plan for the Evolution of the Defense Communications Agency's Office Automation Program, April 1984. Washington, DC.

END

FILMED

1-86

DTIC