

AD 680 111



SP-3243/000/00



Microfilm and Information Retrieval

Arthur Teplitz

29 October 1968

This document is
for public release and its
distribution is unlimited.

D D C
R
JAN 9 1969
B

Reproduced by the
CLEARINGHOUSE
for Federal Scientific & Technical
Information Springfield Va. 22151

SP-3243/000/00

SP *a professional paper*

Microfilm and Information Retrieval

by

Arthur Teplitz

29 October 1968

SYSTEM

DEVELOPMENT

CORPORATION

2500 COLORADO AVE.

SANTA MONICA

CALIFORNIA

90406



29 October 1968

1

SP-3243/000/00

ABSTRACT

This paper discusses the principles of information retrieval, considers subject and classification indexing, and describes elements of coding for manual and machine applications. The implications of information retrieval practices on microfilm information retrieval systems are discussed. Characteristics of information retrieval for manual, semiautomated, and automated systems for aperture cards, microfiche, roll film, and chip systems are considered.

FOREWORD

This paper was prepared for presentation at an American Management Association Seminar on Microfilm Information Retrieval Systems, held at the Ambassador Hotel, Los Angeles, October 21-23, 1968.

As such, it was designed to be delivered in approximately a one-hour session, and to provide a frame of reference for the attendees to the role of microfilm within the information retrieval world.

Consequently, discussions of the elements of the information retrieval characteristics, and of microfilm systems has been simplified to permit some consideration within that limited time period.

The author hopes that the availability of the paper in printed form will provide opportunities for evaluation of the use of microforms for active retrieval applications, and will enable the reader to pursue areas of interest through sources mentioned in the bibliography.

TABLE OF CONTENTS

<u>Section</u>		<u>Page</u>
1.	INTRODUCTION.	1
2.	DEFINITIONS	1
3.	CHARACTERISTICS OF INFORMATION RETRIEVAL.	2
3.1	INDEXING.	4
3.1.1	Subject Terms	4
3.1.2	Classification Systems.	11
3.2	CODING.	11
3.2.1	Manual Codes.	12
3.2.1.1	Derived Coding Methods.	12
3.2.1.2	Assigned Coding Methods	12
3.2.2	Machine Codes	13
4.	CHARACTERISTICS OF MICROFILM RETRIEVAL SYSTEMS.	13
4.1	APERTURE CARD RETRIEVAL SYSTEMS	18
4.1.1	Manual Systems.	18
4.1.2	Semiautomated Systems	19
4.1.3	Fully-Automated Systems	20
4.2	MICROFICHE RETRIEVAL SYSTEMS.	23
4.2.1	Manual Systems.	23
4.2.2	Semiautomated Systems	24
4.2.3	Automated Systems	24
4.3	ROLL AND CARTRIDGE SYSTEMS.	25
4.4	CHIP SYSTEMS.	5
4.5	VIDEOFILE	27
5.	SUMMARY	27
	BIBLIOGRAPHY.	29

LIST OF ILLUSTRATIONS

<u>Figure</u>		<u>Page</u>
1	KWIC Index	5
2	KWOC Index	6
3	Cooccurrence Index	8
4	Uniterm Coordinate Index	9
5	Coordinate Index	10
6	Representative Microforms	16
7	Book Fiche	17

1. INTRODUCTION

The use of microfilm as the storage and dissemination medium for an information retrieval system does not change the essential requirements of that system. The requirements remain constant, regardless of the storage medium. Thus, paper, magnetic tape, or microfilm-based systems have similar characteristics.

It is the purpose of this paper to discuss the characteristics of information retrieval systems, to describe some of the principles of indexing and coding, and to consider the implications of these characteristics and principles as they apply to microfilm-based information systems.

Where applicable, references are made to specific pieces of equipment, by brand name. Such references should not be construed as endorsement of the product, nor of non-recommendation of products designed to do the same task. Rather, the identification should be considered representative of a class of equipment.

A more complete listing of available microfilm equipment may be found in the Annual Reference Guide of Business Automation magazine¹ and other appropriate publications.^{2, 3, 4}

2. DEFINITIONS

- a. Information. (1) Any facts or data which can be used, transferred, or communicated.⁵
(2) The meaningful content of any communication.

- b. Retrieval. (1) Techniques of searching an index file or document collection for information.
- (2) The act of finding again, recovery, retrospective searching and securing of documents.

3. CHARACTERISTICS OF INFORMATION RETRIEVAL*

One approach to evaluation of the characteristics of information retrieval systems is to group these systems into categories and to identify the functions performed within these categories of retrieval systems. Naturally, different authorities describe these groupings and elements in their own unique ways. However, two approaches appear representative of the majority of the authorities cited in this paper.

1. Information retrieval may be classified into three groups⁷ as follows:

- a. Data Base Systems. The most widespread information retrieval systems are those for data base file management, which process records organized into fields, each containing a type of data in the record. Other data base systems include management information systems, and data banks.
- b. Reference Systems. Reference systems require thesauri and process references, rather than data or context. A reference system may be used as a document retrieval system, which may utilize micro-images.
- c. Text Processing Systems. In this system, the full text of the publication is recorded in machine language for subsequent analysis and processing.

*The author particularly commends and recommends Methods of Information Handling, by Charles P. Bourne⁸ and Principles of Automated Information Retrieval, by Walter F. Williams⁹. These books were the primary source for preparation of the material in this section.

2. Information storage and retrieval systems¹⁰ have the following elements:
- a. Input. This involves checking the material, cataloging and indexing the material, and preparing it for input into the storage medium.
 - b. Storage. This involves the conversion and housing of the material itself, and the conversion and housing of the surrogates used in the retrieval functions, such as title, abstract, etc.
 - c. Announcement. This involves processing of the surrogates such that potentially interested users may be made aware of the existence of the material. Note that some announcement systems include the delivery of the material with the announcement, rather than delivery of the substitute for the material.
 - d. Search. This involves matching the requestor's needs with the contents of the stored material, either through manual or automated search and compare of the surrogates, or by search and compare methods through the original material.
 - e. Retrieval. This involves locating and extracting from the store the desired original material often in its stored form, if microfilm, or in its original form, if the stored form is not suitable for subsequent copying or use.
 - f. Delivery. This involves transmission of the original material in usable form. This may be the actual original material, a printed copy of the original, a microfilmed copy, or even a television display of the material.

The task may also be defined as follows:

1. Preparing and using indexes and abstracts.
2. Storing and retrieving documents and information.

The balance of Section 3 is devoted to a discussion of indexing principles, and coding techniques. Abstracts will not be discussed. In Section 4, the characteristics of microfilm information retrieval systems are discussed. Other storage methods are mentioned only briefly.

3.1 INDEXING

Indexing schemes are based on two approaches:

1. Subject terms.
2. Classification systems.

3.1.1 Subject Terms

Subject terms are derived from the document title, the abstract, the full text of the document, or a formally-prepared and maintained list of alphabetic terms. The subject terms may be derived manually, or automatically. Typical techniques include the following:

1. Key-Word-In-Context (KWIC) Index (See Figure 1). Usually, only the title is used and non-significant words such as "and", "the", etc. are not listed as key words. The title is rotated (permuted) as required to list each significant word in alphabetic order within the list. The remaining words in the title are listed in order within the line. KWIC indexes may also be generated solely or partly from key words established by an indexer from the full text.
2. Key-Word-Out-Of-Context (KWOC) Index (See Figure 2). KWOC is essentially the same as KWIC, except that the key word is isolated to the left (or right) of the full title, or summary abstract used in place of the title. Addition of terms to the title (enrichment) is also used to improve the relationship of the subject.

THIRD GENERATION ADP EQUIPMENT OF ALTERNATIVE CHEMICAL S	372
MLA, ERIC, AND THE FUTURE=	572
NUMBER OF PUNCHING ERRORS IN A CO-ORDINATE INDEXING SY	56
RE AND DETECTION OF ERRORS IN PRODUCTION DATA COLLECTIO	612
DESIGN AND INDUSTRY= ESTABLISHING AND USING REFERENCE MA	497
PUBLIC LIBRARY= THE ESTABLISHMENT AND GROWTH OF THE DAT	297
SOCIATIONS PROPOSES ESTABLISHMENT OF A NATIONAL COUNCIL	135
R A TEST PROGRAM TO EVALUATE THE DEMAND SEARCH PERFORMA	468
= A WORLD SYSTEM OF EVALUATED NUMERICAL DATA FOR SCIENC	683
E-ART= CRITERIA FOR EVALUATING THE EFFECTIVENESS OF LIB	705
S= AN EXPERIMENT IN EVALUATING THE QUALITY OF TRANSLATI	105
EARLY INFORMATION EVALUATION=	692
SYSTEM PERFORMANCE EVALUATION: SURVEY AND APPRAISAL=	97
CLASSIFICATION AND EVALUATION= TRACE--TIME-SHARED SOUT	600
SYSTEMS= THE EVALUATION AND ANALYSIS OF INDEXING	290
US= EVALUATION OF A "GROWING" THESAUR	541
ION, OPERATION, AND EVALUATION OF A COMPANY LIBRARY= A	192
EXPERIMENT IN TEXT AND EVALUATION OF AN INFORMATION RETRIE	575
G FILE DEVELOPMENT= EVALUATION OF COORDINATE INDEX SYST	339
AND SERVICES= EVALUATION OF INFORMATION SYSTEMS A	535
ORY APPROACH TO THE EVALUATION OF INFORMATION RETRIEVAL	217
FOR MANAGEMENT - AN EVALUATION OF MANAGERIAL REQUIREMEN	229
BIBLIOGRAPHERS= AN EVALUATION OF REGULARS OUTPUT: DEMAN	285
S= THE QUANTITATIVE EVALUATION OF REGIONAL UNION CATALO	90
DATA -- A CRITICAL EVALUATION OF RESULTS IN 100 CASES	347
TATION INDEXING AND EVALUATION OF SCIENTIFIC PAPERS= CI	420
LIST FOR REVIEW AND EVALUATION OF TECHNICAL LIBRARIES=	629
E THROUGH SIMULATED EVOLUTION= ARTIFICIAL INTELLIGENC	193
TE AN IBM-7090/1604 EXAMPLE= TRANSLATION BY IACT -- WI	498
INFORMATION EXCHANGE=	304
ORK FOR COOPERATIVE EXCHANGE) -- BOOK CATALOG PRODUCTI	573
MENT (INFORMATION EXCHANGE GROUP)= DEATH OF AN EXPERI	236
MENT OF INFORMATION EXCHANGE GROUPS= INTERNATIONAL STAT	653
ADMINISTRATOR= THE EXECUTIVE AND THE INFORMATION EXPLO	334
DERM LITERATURE= AN EXEMPLARY THEORY OF COMMUNICATION I	694
HE SEED FOR CATALOG EXPANSION= PREDICTING T	273
MEDIANS OPERATING EXPERIENCE: ADDENDUM=	546
SSITY ARCHIVES: THE EXPERIENCE OF ONE INSTITUTION= COLL	448
ENCE CENTER= EARLY EXPERIENCE WITH A TECHNICAL CORRESP	120
ATION SYSTEM= ISI'S EXPERIENCE WITH ASCA -- A SELECTIVE	204
REVIEWS: A COMPUTER EXPERIMENT= AN INDEX TO MEDICAL BOO	31
TENCY IN TEXT -- AN EXPERIMENT AND SOME SUGGESTIONS= SE	713
L TRANSLATIONS= AN EXPERIMENT IN EVALUATING THE QUALIT	105
L PROCEDURES FOR AN EXPERIMENT IN TEXT AND EVALUATION O	575
ND PROJECT NARC; AN EXPERIMENT IN THE DISTRIBUTION OF N	36
GROUP)= DEATH OF AN EXPERIMENT (INFORMATION EXCHANGE	236

Figure 1. KWIC Index

NON-IRRADIATED	ABSORPTION OF D-GLUCOSE BY SEGMENTS OF INTESTS FROM ACTIVE AND HIBERNATING, IRRADIATED AND NON-IRRADIATED GRAY SQUIRRELS, SIZELLUS 312 DECEMLINEATUS NASA 603-110021K) 82.60 0726		
NON-ISOTHERMAL	CORRELATIONS IN A NON-ISOTHERMAL PLASMA AD-290 0531K) 81.10 0196	NUCLEAR	ETIC BLACKOUT FOLLOWING A HIGH ALTITUDE NUCLEAR DETONATION AD-291 1411K) 88.60 0372
NON-LINEAR	INVESTIGATION OF MICROWAVE NON-LINEAR EFFECTS UTILIZING FERROMAGNETIC MATERIALS AD-290 0721K) 82.60 0497	NUCLEAR	ACCURATE NUCLEAR FUEL BURNUP ANALYSES GEAP-40521K) 81.60 0362
NON-METALLIC	BIBLIOGRAPHY AND TABULATION OF DAMPING PROPERTIES OF NON-METALLIC MATERIALS AD-289 0561K) 83.00 0502	NUCLEAR	APPLICATION OF NUCLEAR POWER SUPPLIES TO SPACE SYSTEMS TID-173061K) 81.60 0741
NON-MILITARY	NOYES ON NON-MILITARY MEASURES IN CONTROL OF INSURGENCY AD-290 0371K) 81.60 0496	NUCLEAR	CAROLINAS-VIRGINIA NUCLEAR POWER ASSOCIATES, INC., RESEARCH AND DEVELOPMENT PROGRAM QUARTERLY PROGRESS REPORT FOR THE PERIOD APRIL - JUNE 1967 CVNA-1501K) 86.60 0339
NON-MOVING	JUDGMENTS OF VISUAL VELOCITY AS A FUNCTION OF THE LENGTH OF OBSERVATION TIME OF MOVING OR NON-MOVING STIMULI PB 162 5491K) 81.60 0145	NUCLEAR	COMPUTER PROGRAMS FOR OPTIMUM START-UP OF NUCLEAR PROPULSION SYSTEMS TID-167301K) 81.10 0712
NON-RELATIVISTIC	TABLES OF NON-RELATIVISTIC ELECTRON TRAJECTORIES FOR FIELD EMISSION CATHODES AD-290 0961K) 81.60 0239	NUCLEAR	DOSE-RATE-DISTANCE CURVES FOR CLOSE-IN FALLOUT FOR LOW YIELD LAND-SURFACE NUCLEAR DETONATIONS PB 162 5491K) 81.60 0573
NON-SIMILAR	NON-SIMILAR NUMERICAL METHODS OF SOLUTION FOR ELECTRODE BOUNDARY LAYERS IN A CROSSED FIELD ACCELERATOR AD-290 3251K) 85.60 0185	NUCLEAR	EXTRUDED CERAMIC NUCLEAR FUEL DEVELOPMENT PROGRAM ANP-625501K) 84.60 0092
NONDESTRUCTIVE	NONDESTRUCTIVE SYSTEM FOR INSPECTION OF FIBER GLASS-REINFORCED PLASTIC MISSILE CASES AD-289 0251K) 81.60 0632	NUCLEAR	FEASIBILITY DETERMINATION OF A NUCLEAR THERMOELECTRIC SPACE POWER PLANT AD-290 0691K) 82.60 0031
NONDESTRUCTIVE	X-RAY IMAGE SYSTEM FOR NONDESTRUCTIVE TESTING OF SOLID PROPELLANT MISSILE CASE WALLS AND WELDMENTS AD-289 0711K) 83.60 0637	NUCLEAR	HIGH-ENERGY NUCLEAR PHYSICS RESEARCH PROGRAM AD-291 1401K) 81.60 0374
NONDISSIPATIVE	MAGNETOHYDRODYNAMIC STABILITY OF VORTEX FLOW - A NONDISSIPATIVE, INCOMPRESSIBLE ANALYSIS ORNL-TM-4021K) 83.60 0615	NUCLEAR	HIGH-ENERGY NUCLEAR REACTIONS OF NIOBIUM WITH INCIDENT PROTONS AND HELIUM IONS UCRL-10461K) 82.25 0222
NONEQUILIBRIUM	SCALE EFFECTS FOR NONEQUILIBRIUM CONVECTIVE HEAT TRANSFER WITH SIMULTANEOUS GAS PHASE AND SURFACE CHEMICAL REACTIONS. APPLICATION TO HYPERSONIC FLIGHT AT HIGH ALTITUDES AD-291 0321K) 81.60 0625	NUCLEAR	INVESTIGATIONS ON THE DIRECT CONVERSION OF NUCLEAR FISSION ENERGY TO ELECTRICAL ENERGY IN A PLASMA DIODE AD-290 7271K) 89.60 0385
NONLINEAR	APPLICATION OF VARIATIONAL EQUATION OF MOTION TO THE NONLINEAR VIBRATION ANALYSIS OF HOMOGENEOUS AND LAYERED PLATES AND SHELLS AD-289 0881K) 82.60 0667	NUCLEAR	NUCLEAR SUPERHEAT DEVELOPMENT PROGRAM GNEC-2541K) 81.60 0286
NONLINEAR	EXTENSIONS IN THE SYNTHESIS OF TIME OPTIMAL OR BANG-BANG NONLINEAR CONTROL SYSTEMS. PART I. THE SYNTHESIS OF QUASI-STATIONARY OPTIMUM NONLINEAR CONTROL SYSTEMS PB 162 5471K) 84.60 0235	NUCLEAR	PRODUCTION OF TRITIUM BY CONTAINED NUCLEAR EXPLOSIONS IN SALT. 1. LABORATORY STUDIES OF ISOTOPIC EXCHANGE OF TRITIUM IN THE HYDROGEN-WATER SYSTEM ORNL-3334K) 83.60 0617
NONLINEAR	EXTENSIONS IN THE SYNTHESIS OF TIME OPTIMAL OR BANG-BANG NONLINEAR CONTROL SYSTEMS. PART I. THE SYNTHESIS OF QUASI-STATIONARY OPTIMUM NONLINEAR CONTROL SYSTEMS PB 162 5471K) 84.60 0235	NUCLEAR	STRIKING EFFECT OF NUCLEAR EXPLOSION AD-290 8241K) 82.10 0083
NONLINEAR	NONLINEAR FLEXURAL VIBRATIONS OF SANDWICH PLATES AD-289 0711K) 82.60 0669	NUCLEAR	THE NUCLEAR PROPERTIES OF RHENIUM AD-291 1801K) 81.60 0310
NONLINEAR	OPTIMUM NONLINEAR CONTROL FOR ARBITRARY DISTURBANCES NASA 602-158901K) 82.60 0682	NUCLEAR	VARIATIONS IN THE TOTAL ELECTRON CONTENT OF THE ATMOSPHERE AFTER THE HIGH ALTITUDE NUCLEAR EXPLOSION NASA 603-104861K) 81.10 0142
NONRECURRENT	A TECHNIQUE FOR NARROW-BAND TELEMETRY OF NONRECURRENT PULSES AD-290 6971K) 82.60 0577	NUCLEAR	430A MARITIME NUCLEAR STEAM GENERATOR GEMP-1601K) 80.10 0349
NONUNIFORM	ELECTROMAGNETIC SCATTERING FROM A SPHERICAL NONUNIFORM MEDIUM. PART II. THE RADAR CROSS SECTION OF A FLARE AD-289 6151K) 82.60 0747	NULL-ZONE	THE ESTIMATION PROBLEM IN NULL-ZONE RECEIVING FEEDBACK SYSTEMS AD-290 325 1) 81.10 0599
NONUNIFORM	ELECTROMAGNETIC SCATTERING FROM ASPHERICAL NONUNIFORM MEDIUM. PART I. GENERAL THEORY AD-289 6141K) 82.60 0748	NUMBERS	FUNDAMENTAL SOLUTION TO THE DIFFUSION BOUNDARY LAYER EQUATION FOR NEARLY SEPARATED FLOW OVER SOLID SURFACES AT VERY LARGE PRANDTL NUMBERS AD-291 0311K) 82.60 0223
NORMAL	PROBABILITY INTEGRALS OF MULTIVARIATE NORMAL AND MULTIVARIATE-T AD-290 7401K) 80.60 0760	NUMBERS	LOCAL PRESSURE DISTRIBUTION ON A BLUNT DELTA WING FOR ANGLES OF ATTACK UP TO 35-DEGREES AT MACH NUMBERS OF 3.4 AND 4.7 AD-291 0311K) 82.60 0223
NORMAL	RESONANCE ABSORPTION OF GAMMA-RAYS IN NORMAL AND SUPERCONDUCTING TIN AD-289 0441K) 83.60 0826	NUMERICAL	A MAINTENANCE PROGRAM FOR NUMERICAL CONTROL SYSTEMS ON MACHINE TOOLS TID-173761K) 82.60 0609
NORMS	NORMS FOR ARTIFICIAL LIGHTING AD-290 5551K) 81.10 0734	NUMERICAL	A PRIORI BOUNDS ON THE DISCRETIZATION ERROR IN THE NUMERICAL SOLUTION OF THE DIRICHLET PROBLEM AD-290 3221K) 84.60 0664
NORTH	FACTORS INFLUENCING VASCULAR PLANT DONATION IN NORTH CAROLINA SALTWATERS AD-290 9381K) 87.60 0603	NUMERICAL	NON-SIMILAR NUMERICAL METHODS OF SOLUTION FOR ELECTRODE BOUNDARY LAYERS IN A CROSSED FIELD ACCELERATOR AD-290 5251K) 85.60 0185
NORTH	SONAR STUDIES OF THE DEEP SCATTERING LAYER IN THE NORTH PACIFIC PB 167 4271K) 82.60 0587	NUSTAGMUS	MANIPULATION OF AEROSOLS AND ITS EFFECTS ON HUMAN VESTIBULAR NYSTAGMUS INDUCED BY CALORIC IRRADIATION AND ANGULAR ACCELERATIONS AD-290 3481K) 81.60 0252
NORTH	THE DEVELOPMENT OF RESCUE AND SURVIVAL TECHNIQUES IN THE NORTH AMERICAN ARCTIC PB 162 4101K) 81.20 0085	OAK	A SAFETY REVIEW OF THE OAK RIDGE CRITICAL EXPERIMENTS FACILITY ORNL-TM-3491K) 85.60 0617
NOSE	THE FLORA OF HEALTHY DOGS. I. BACTERIA AND FUNGI OF THE NOSE, THROAT, AND LOWER INTESTINE (P-21K) 82.60 0458	OBJECTS	DRAW OF OBJECTS IN PARTICLE - LOADED AIR FLOW PHASE IV. BLUNT BODIES AND COMPRESSIBILITY EFFECTS AD-291 1781K) 80.60 0752
NOZZLE	FABRICATION OF PYROLYTIC GRAPHITE ROCKET NOZZLE COMPONENTS PB 162 3711K) 81.10 0351	OBSERVATORY	TONTO FOREST SEISMOLOGICAL OBSERVATORY AD-291 1401K) 83.60 0815
NOZZLE	FABRICATION OF PYROLYTIC GRAPHITE ROCKET NOZZLE COMPONENTS PB 162 3701K) 81.10 0353	OCEAN	A SAMPLE TEST EXPOSURE TO ERASIVE CORROSION AND FOULING OF EQUIPMENT INSTALLED IN THE DEEP OCEAN AD-291 0491K) 81.60 0582
NOZZLE	FABRICATION OF PYROLYTIC GRAPHITE ROCKET NOZZLE COMPONENTS PB 162 3721K) 82.60 0352	OCEANOGRAPHIC	OCEANOGRAPHIC CRUISE TO THE BERING AND CHUKCHI SEAS, SUMMER 1949. PART I. SEA FLOOR STUDIES PB 162 4261K) 82.60 0585
NOZZLE	THIRD SYMPOSIUM ON ADVANCED PROPULSION CONCEPTS SPONSORED BY UNITED STATES AIR FORCE OFFICE OF SCIENTIFIC RESEARCH AND THE GENERAL ELECTRIC COMPANY FLIGHT PROPULSION DIVISION CINCINNATI, OHIO OCTOBER 2-4, 1962. PLASMA FLOW IN A MAGNETIC ARC NOZZLE AD-290 0071K) 82.60 0147	OCEANOGRAPHIC	OCEANOGRAPHIC AND UNDERWATER ACOUSTICS RESEARCH AD-290 2571K) 82.60 0049
NOZZLES	HEAT TRANSFER AND PARTICLE TRAJECTORIES IN SLUG-ROCKET NOZZLES AD-289 0711K) 85.60 0020	OCEANOGRAPHIC	OCEANOGRAPHIC CRUISE TO THE BERING AND CHUKCHI SEAS, SUMMER 1949. PART IV. PHYSICAL OCEANOGRAPHIC STUDIES. VOL. 1. DESCRIPTIVE REPORT PB 162 420-11K) 83.60 0564
		OCEANOGRAPHIC	OCEANOGRAPHIC CRUISE TO THE BERING AND CHUKCHI SEAS, SUMMER 1949. PART IV. PHYSICAL OCEANOGRAPHIC STUDIES. VOL. 2. DATA REPORT PB 162 420-21K) 84.60 0566

Figure 2. KWOC Index

3. Clue-Word Index (See Figure 3). Each index line is essentially a summary abstract and is repeated in many locations throughout the index. Descriptors are established by the indexer, and the abstract is repeated as often as is deemed appropriate. The Co-occurrence Index¹¹ shown in Figure 3 is one approach to clue word indexing. It utilizes a controlled vocabulary of subject terms and descriptors.
4. Uniterm Coordinate Indexing (See Figure 4). This is an alphabetic file of single-word subject terms assigned to each document. Any number of terms may be used (theoretically) to describe a document. That document may be found by a subject search (not serial) of the alphabetic listing. Each document is listed under the pertinent terms. All terms are considered equivalent in value...there is no subordination in relationships.
5. Coordinate Indexing (See Figure 5). Coordinate Indexing in any of its several forms is the most common subject indexing technique. The approach is based on the Uniterm concept of unique identifiers for each document, but frequently utilizes such concepts as roles, links, modifiers, and subject term enrichment. Use of these elements tends to reduce false drops, to provide for more uniformity in terms, and to enhance the effectiveness of the index...they say.

Coordinate indexing tends to have a relatively smaller size index in comparison to collections than many other indexing approaches. A lightly-indexed file may use an average of 5 terms per document, up to 15 terms. A heavily-indexed file, utilizing links, roles, and modifiers may use 65-100 entries, or even higher. Typical relationships of terms to collections are as follows:

EQUIP(MENT) (SEE ALSO 'FAC' AND SPECIFIC TERMS).

BARRINGTON/KARDON, COMP+, EQUIP, -SDC SM BY 4 NOV/67 09 05A
 BARRINGTON/KARDON REQ, SDC AID, EQUIP, INSTAL/ 67 09 05A
 HASKIN/SEAMO, MAN TROUB=, *SKED SLIP, EQUIP, -SDC. 67 09 07B ←

GAZOTA KNOB(S).

RICH, -HASKIN/ BORAD, RFP RE, GAZOTA KNOB, -SDC/ 67 08 31A
 HASKIN/MEET 4-5 SEP RE, SDC, PROP RE, GAZOTA KNOB/ 67 09 07B ←
 HASKIN/BORAD+, SEAMO OK, GAZOTA KNOB, DESIGN/ 67 09 07B ←
 HASKIN/BORAD JOB, GAZOTA KNOB, ACCEP, TEST PROCED/67 09 07B ←

HASKIN R SDC.

GOLD/ MEET 12 AUG RE, COST REV, HASKIN=, SDC REP/ 67 08 20B
 HASKIN, -RICH/MEET 4-5 SEP, BORAD+, SEAMO+, HASKIN/ 67 09 07B ←

MAN(NING) OR MANPOWER (SEE ALSO 'PERS(ONNEL)').

ARMOR, -OSWELL/KARDON REQ, SDC CONSID, *SITE, MAN/ 67 09 01A
 HASKIN/BORAD NT, MAN+, COST EST BY 1 DEC/ 67 09 07B ←
 HASKIN/SFAMO, MAN TROUB=, *SKED SLIP, EQUIP, -SDC. 67 09 07B ←

MEET(ING).

GOLD/ MEET 12 AUG, BASCO+, SDC RE, ACCEP+, COST/ 67 08 20B
 GOLD/ MEET 12 AUG RE, COST REV, HASKIN=, SDC REP/ 67 08 20B
 HASKIN, -RICH/MEET 4-5 SEP, BORAD+, SEAMO+, HASKIN/ 67 09 07B ←
 HASKIN/MEET 4-5 SEP RE, SDC, PROP RE, GAZOTA KNOB/ 67 09 07B ←

The first indexing line for the source item also appears under the recipient's name in the Author-Recipient Index, as shown:

RICH I SDC.

RICH, -OSWELL/ BORAD, RFP RE, GAZOTA KNOB, -SDC. 67 08 31A
 HASKIN, -RICH/MEET 4-5 SEP, BORAD+, SEAMO+ HASKIN/ 67 09 07B ←

Figure 3. Cooccurrence Index

MISSILES

00020	01351	00652	00653	00234	01445	01076	01117	00658	00149
01080	02901	02882	02443	00824	01855	01886	00217	00688	00659
01240	02651	05912	02553	01074	01995	02496	01087	01758	01739
02280	03291	06912	02573	01274	03335	02546	01247	01008	02199
03900	05451	07972	03073	02494	03535	02566	01797	02296	02359
04240	04231	11332	04063	02614	04685	02646	01997	02488	03189
04250	04451		05613	02714	04895	03536	02627	02598	04229
04890	04941		05893	03154	05235	04136	04137	03158	04239
05910	04991		05913	03454	05285	05056	04237	03205	04969
07600	05301		06583	04054	06265	06536	04517	04228	05249
07780	05921		06763	04244	06405	06586	04887	04338	05889
11160	06541		06843	05414	06783	07506	05107	04778	05909
11180	07141		07053	05614	07335	08968	05217	04948	06489
11810	07331		07143	05914	07375	12526	05237	04968	06559
12220	07501		07971	06764	09845		05337	05248	06569
	07971		08103	06914	10125		05367	05908	10209
	08031		08283	09864			07827	06568	
	08131		10243	10104			10317	06588	
	09151		10353	10384			12127	06598	
	10211		10443	12734			12317	10428	
	10381		11353				12377	12308	
	10441		11633					12348	
	11611		12723					12368	
	12301								

POISEVILLE FLOW									
	07191			01283					08829
	09871								
POISSON									
	05361	02782	02783	02784					10199
		05362							
POISSONS RAYIO									
	00380								
POLAR								06184	11675
POLAR ORBITS									
	10291							06623	
POLAR REGIONS									
	04840	03311	00302	01343	02144	00165		02437	01468
		03671		07443	07834			04647	12008
		10161			09264			06019	07517
POLARIMETERS									
	02240								
POLARIS									
	02140	03451	00652	01193	05874	05235	05896	00568	02009
				02553			07506	09138	10489
				06493			09026	12308	

- 648. Personnel management poses one of most critical problems. p.26-. PERSONNEL - MANAGEMENT
- 649. Metals fastening know-how brings helicopter rotor blade-making. p.32-. HELICOPTERS - ROTORS - BLADES - MANUFACTURE

MISSILES & ROCKETS Jan. 13, 1964

- 650. Gemini model readied for water recovery testing. p.16-. GEMINI - MODELS - RECOVERY - TESTING
- 651. Safe design of reactors is indicated. p.17-. NUCLEAR - REACTORS - REENTRY - SAFETY
- 652. Proposed fourth-generation Polaris would use state-of-art technology. Lindsey, R. p.18-21 POLARIS - MISSILES - REENTRY
- 653. Astrolog - current status of U.S. missile and space programs. p.25-31 DIRECTORIES - CONTRACTS, CONTRACTORS, CONTRACTING - MISSILES - ARTIFICIAL SATELLITES - SPACE VEHICLES
- 654. Gemini photos may assist in re-entry wake research. Hawkes, R. p.34-. GEMINI - REENTRY - WAKES - PHOTOGRAPHY
- 655. Space tool for Apollo can deliver 50 ft.-lbs. of work. p.37-. TOOLS - SPACE FLIGHT - APOLLO - MANNED - SPACE VEHICLES - REPAIR - TORQUE - LIFE SUPPORT SYSTEMS
- 656. Contractor gets steady check on profits with new PERT-type system. Beller, W. p.38-41 PERT - MANAGEMENT SYSTEMS - PROFITS - CONTRACTS, CONTRACTORS, CONTRACTING

NOTE: The example shown here is taken from the Pacific Aerospace Library's Dual Dictionary. PAL is no longer using the dual dictionary, but it remains an outstanding example of the Uniterm approach.

Figure 4. Uniterm Coordinate Index

theory REPT.-68/14	c24 N68-29843	Supersonic transport fatigue design and test program	c32 N68-29913
FAIL-SAFE SYSTEMS		Fatigue tests and analysis of piston Provost wings and progressed loading	
Fokker F-28 transport aircraft fatigue test plan to demonstrate user specifications and prove safe life and fail-safe requirements	c32 N68-29699	S+T-MEMO-5/67	c32 N68-30298
FAILURE		Fatigue tests results and analysis of wing loading in ascending and descending order for piston Provost wings	
Stabilization and control system sensitivity to power-off failure rate studied by simulated missions using block power switching		S+T-MEMO-1/68	32 N68-30299
NASA-CR-6085	c31 N68-30182	Structural acoustics problems in Concorde supersonic civil aircraft development	
FAR INFRARED RADIATION		SST/B745-7/2966	c02 N68-30300
Electrical properties of materials in far infrared region - vacuum monochromator development and performance testing		FEASIBILITY	
AEDSR-68-0465	c14 N68-29814	Feasibility of power by nuclear fusion	
PARADAY EF CT		ORNL-TM-2204	c25 N68-30162
Atlas of lunar data with plots of ionospheric Faraday rotation and electron content versus time		FEEDBACK CONTROL	
AD-669634	c50 N68-30853	Void fraction distributions in vertical concurrent gas-liquid flow in closed loop system	
FAST NEUTRONS		Performance testing of NERVA XE-1 control drum actuator to determine its ability to function in closed loop positioning system	
Experiment for determining solar neutrons in atmosphere during solar eclipse		NASA-TM-X-52465	c22 N68-29973
AC-T-68-1518	c29 N68-30710	Energy flow equilibrium, heat transfer, and closed loop feedback in hot-wire anemometers	
FAST NUCLEAR REACTORS		TR-2	c14 N68-30054
Neutronic effects of yttrium hydride moderator insertion in fast spectrum reactors		Design of nonlinear control systems using state variable feedback	
NASA-TM-X-1514	c22 N68-29950	Signal transmission in linear or nonlinear systems with and without feedback related to radio engineering, electronics, and automation	
Linear analytical dynamic model developed for steam cooled fast power reactors		FTD-HT-67-191	c07 N68-30363
AE-316	c22 N68-30215	Cybernetics textbook expounding theory of adaptive control systems combining open loop disturbance coupling and feedback control	
Experimental information from dilute plutonium fueled critical assemblies used to improve fast reactor design data and calculation methods		FTD-HT-23-578-67	c10 N68-30562
GEAP-5271	c22 N68-30685	FERMI SURFACES	
Neutron dosimetry and spectra used to determine effectiveness of fast nuclear reactor neutron irradiations		Two spin deviate problem for exact analysis of interacting Fermi gas	
BNWL-SA-1488	c22 N68-30703		c24 N68-29642
Electron probe microanalysis of radioactive materials used in fast breeder reactors		FERRITES	
GEAP-5344	c11 N68-30714	Theoretical analysis of digital circuits based on principles of current distribution and methods for circuit calculations of transistorized and ferrite-core containing units	
Fast Flux Test Facility for liquid metal fast breeder reactor materials and fuels		FTD-HT-24-123-67	c10 N68-29946
BNWL-SA-1750	c22 N68-30722	Feasibility study of random access laminated ferrite memory system for spacecraft use	
Reactivity and neutron spectra in fast nuclear reactors, and sub-critical time-of-flight spectrum facility		NASA-CR-66636	c08 N68-30088
GA-8377	c22 N68-30725	Operation of external access ferrite rectangular hysteresis loop storage devices	
Neutronic calculations of fuel and poison drum control of refractory metal, fast spectrum space power reactors		FTD-HT-23-915-67	c08 N68-30106
NASA-TN-D-4709	c22 N68-30750	Relating ferrite decomposition and associated precipitates in stainless steel to its corrosion resistance	
Performance potential of U-Pu ceramic fuels for fast liquid metal cooled reactors		FTD-HT-67-223	c17 N68-30484
BNWL-SA-1550	c22 N68-30785	FERROUS	
FATIGUE (MATERIALS)		X ray absorption spectra and calculated charges of iron atoms in polyferrocenes	
Electron fractography used to study fatigue crack propagation resulting from cyclic crack extension by alternating shear		FTD-HT-66-112	c06 N68-30178
DI-62-0708	c32 N68-29710	FERROELECTRICITY	
Supersonic transport fatigue design and test program		Electrical resistivity measures in charge transfer analysis in lead zirconate titanate	
REPT.-2	c26 N68-29564	Delayed coincidence Mossbauer effect studied for iron compounds	
ARC0-6338.10-P	c26 N68-30849	FERRONAGNETIC MATERIALS	
FERRONAGNETIC MATERIALS		Nonuniform magnetic field effect on ferromagnetic colloidal solutions	
NASA-TN-D-4676	c06 N68-30644	FERRONAGNETISM	
FATIGUE LIFE		Interelectron interactions, energy bands, and spin angular momentum studied to explain ferromagnetism in transition metals	
Probability distribution functions used to find fatigue life, and its variability, of Al alloys subjected to cantilever bending			c20 N68-29797
FTD-HT-66-79	c17 N68-30135	FERTILIZATION	
Neutron irradiation effects on tensile, impact, and fatigue properties of steels and welds in Elk River Reactor /ERR/ pressure vessel		Comparative morphological and teratological study on well-formed frog eggs	
SWRI-1228-P-9-15	c22 N68-30302	NASA-TT-F-11685	c04 N68-30021
Qualitative prediction methods for reliability, wear, and service life of machine tools with fatigue as life limiting factor		Fertilization of overmature frog eggs	
FTD-HT-67-22	c15 N68-30446	NASA-TT-F-11763	c04 N68-30397
		Sexual differentiation of daphnids and frogs	

Figure 5. Coordinate Index

<u>Collection Size</u>	<u>Terms</u>
100	500
1,000	2,000
5,000	5,000
25,000	6,500
100,000	10,000

3.1.2 Classification Systems

Classification systems employ the principle of classes of information, which permits hierarchical relationships to be shown. Typical of the classification schemes are the familiar Library of Congress, Dewey Decimal, and Universal Decimal library systems.

Classification systems are the most complex to establish, and are difficult to automate. They do provide a means of handling very large collections.

Some of the better-known systems and associated names include:

1. Dewey Decimal System - Melvin Dewey
2. Library of Congress System - Charles Cutter
3. Colon Classification System - S. R. Ranganathan and T. Tyaganatarajan
4. Bliss Classification System - H. E. Bliss
5. Faceted Classification System - B. C. Vickery and D. J. Foskett
6. Universal Decimal Classification System - Jean Perrault

It is of interest in these lists to see some of the lineal relationships. For example, the UDC system is a direct descendent of the Dewey system, Foskett took on the Colon Classification scheme, and made modifications to it.

3.2 CODING

Coding is the technique by which the amount of data to be processed by hand or by machine is reduced to tolerable limits. The coding methods employed must be compatible with the storage and retrieval methods used. Many lengthy studies have been done on coding requirements and methods. These studies are beyond the scope of this discussion. However, some of the coding techniques developed as a result of the study are listed in paragraph 3.2.1.

Machine codes are restricted by the equipment used in entering and transmitting the information for machine processing. Manual coding methods for reducing the amount of data are applicable to machine methods. The machine codes, which are based on the available bits in which to mark the existence (or non-existence) of a condition, are carriers of the manual codes. Machine codes are listed in paragraph 3.2.2.

3.2.1 Manual Codes

Manual codes are used for abbreviations of names and English-language words. These codes may be either of the following types:

1. Derived. Derived codes are obtained from manipulation of the term to be encoded, so as to reduce the numbers of letters or terms required to express the full range of values of the terms. For example: Annapolis may be expressed as Annpls, Anpls, Anna.
2. Assigned. Assigned codes have meaning only within a particular frame of reference, and are completely arbitrary as to meaning. For example: 1 = dollars, 2 = lira, 3 = pounds.

3.2.1.1 Derived Coding Methods

1. Truncation
2. Elimination of Vowels
3. Selective Dropout
4. Check Digit
5. Arithmetic Manipulation
6. Others

3.2.1.2 Assigned Coding Methods

1. Arbitrary
2. Notational
3. Prime Number
4. Others

3.2.2 Machine Codes

1. 5-Bit Code (for example, Baudot) - 5 bits. Provides for 32 characters, expandable to 58 characters. Used for teletype, etc.
2. Binary-Coded Decimal (BCD) Interchange Code - 6 bits. Provides for up to 64 characters. Used in machine data processing, facsimile transmission, etc.
3. United States of America Standard Code for Information Interchange (USASCII) - 7 bits. Provides for up to 128 characters.
4. Extended Binary-Coded Decimal (EBCD) Interchange Code - 8 bits. Provides for up to 256 characters.
5. Others

4. CHARACTERISTICS OF MICROFILM RETRIEVAL SYSTEMS

Microfilm retrieval systems have the following elements in common:

1. The individual pages have been photographically reduced.
2. The microform is housed in a file or other container.
3. File numbers, indexing, or coding information is marked on the microform in normal size, when retrieval is essentially manual, or in reduced size when retrieval is automatic.

Individual systems may be measured in terms of effectiveness, based on the following considerations:

1. Completeness of file.
2. Availability of information to multiple users.
3. Desirability of centralized files and/or satellite files.
4. Time required to obtain information from files.
5. Security of information, including file integrity.
6. Cost of system, including indexing, coding, storing, retrieving, and distributing.
7. Reliability of system...accuracy of searching, completeness of retrieval, usefulness of deliverable documents.
8. Ease of use.
9. Cost of alternative systems vs. cost of this system.
10. Growth capabilities...increased volume, increased numbers of users, technological changes.

The relationship of the microform to the document must also be considered.

Some of the elements of particular significance include:

1. Utilization - 1 microform to one document unit.
2. Color requirements - black and white, other.
3. Size constraints...reduction requirements greater than resolving capacity of system.

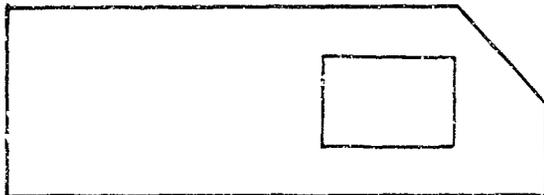
4. Number of copies required.
5. Equipment requirements.

The following microforms (See Figure 6) may be considered for use in information retrieval systems:

1. Aperture Cards. These are unitized tab cards, containing both machine codes and microfilm images. Normally, a single frame of 35mm film is used, but many variations are used.
2. Microfiche. Microfiche are film sheets, usually 105 x 148 millimeters (4 x 6 inches), with a normal reading title block, and rows of reduced page images. Several reduction ratios are commonly used, including
 - a. COSATI - 20:1 with 60 images
 - b. NMA - 24:1 with 98 images.
 - c. Ultrafiche - 140:1 with approximately 1,500 pages.

The author has advocated the use of microfiche for book applications.¹² Book Fiche (also called Library Fiche) uses a single-stage reduction ratio of 50:1, which provides a capacity of 390 pages per fiche (See Figure 7), in addition to the title block, as used on COSATI fiche. This is not the place to argue the merits of Book Fiche. However, the significant advantages of the microform should at least be mentioned:

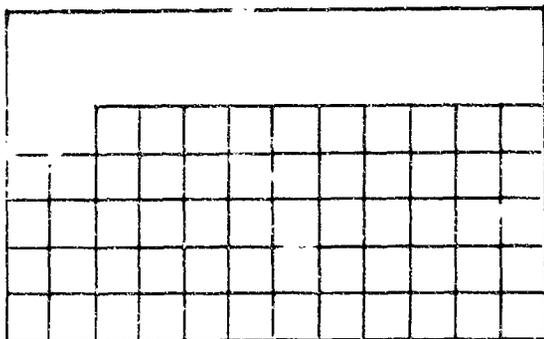
- a. Libraries can function as distribution libraries, rather than circulating libraries.



APERTURE
CARD

REDUCTION IMAGES

16:1 1



COSATI FICHE

20:1 60

NMA COMMERCIAL

24:1 98

BOOK FICHE

50:1 390

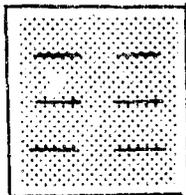
ULTRA FICHE

120:1 2,100
150:1 3,200



ROLL & CARTRIDGE
(16mm or 35mm)

24:1 2,000



FILM CHIP

35mm Chip
to
70mm Chip

10:1 1

40:1 1

Figure 6. Representative Microforms

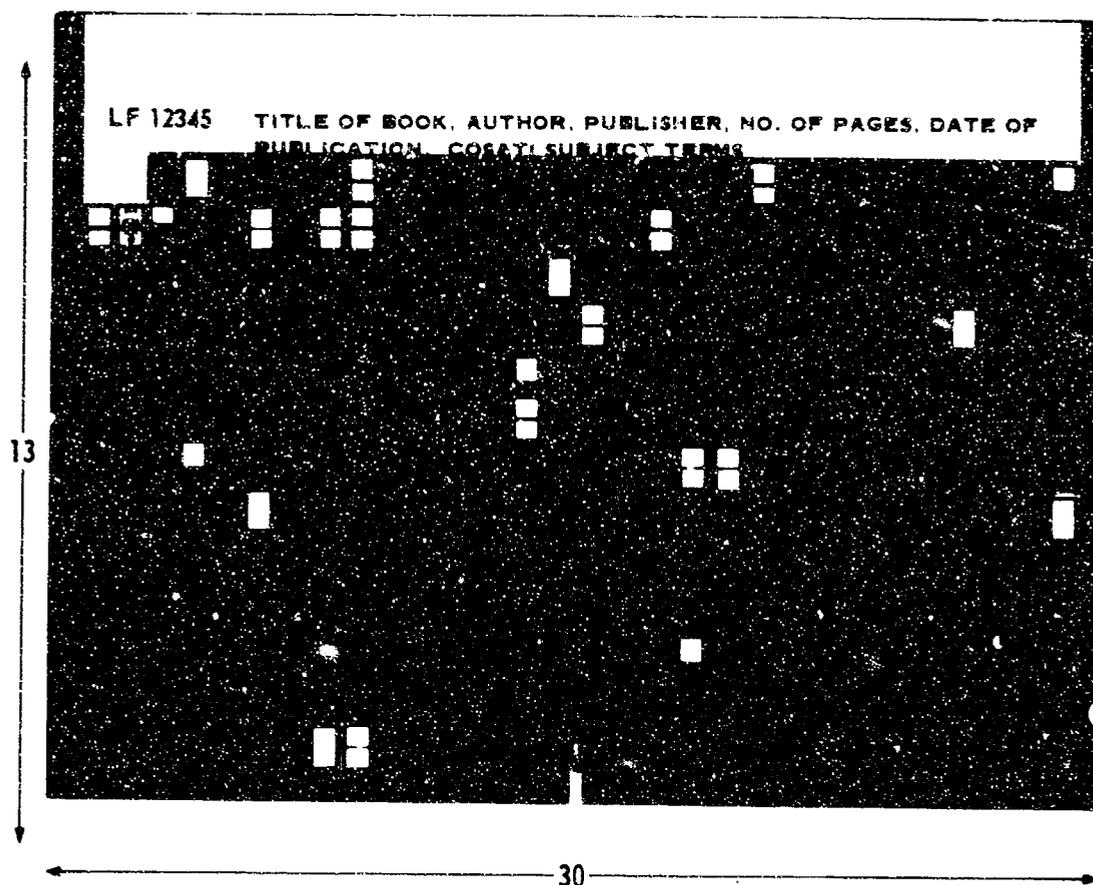


Figure 7. Book Fiche

- b. Books, in fiche form, could be available at a cost of 15 to 20 cents, including royalties, instead of present costs ranging upward from an average of \$7.50 for technical books.
 - c. Retrieval problems are reduced because of the one-to-one relationship of Book Fiche to the book.
3. Roll and Cartridge Film. Roll film is usually 35mm film, cartridge film is usually 16mm film. Roll film often has a reduction ratio of 16:1, providing about 1,800 pages per 100 foot roll. Cartridge film uses a reduction ratio of 28:1, with 2,000 to 3,000

images per roll. Other film and reduction sizes are common, as well, including 70mm, and 105mm at reductions of 5-8:1 and 8mm film at a reduction ratio of 38:1.

4. Chip Systems. Film chips are often used in highly mechanized systems, and are usually 35mm or larger in size. These are unique systems, with reduction ratios, coding, and capacities adapted to the individual systems.
5. Others. Other microimaging systems are found, in which the characteristics are those of microfilm, but the carrier is unique. Videotape, using magnetic tape, rather than microfilm is one example. Filmstrips, as used in the Microstrip system of Eastman Kodak may be considered a special system, or may be considered as a variation of roll film or microfiche.

4.1 APERTURE CARD RETRIEVAL SYSTEMS

4.1.1 Manual Systems

Several possibilities exist for retrieval of aperture cards, based on straightforward retrieval methods, and not using automated handling methods for the aperture card itself. These include:

1. Sequential Filing - Hand-Posted Accession Numbers. The aperture cards are stored in a file, with dividers. Each card has the appropriate accession number written on it, with other identifying information, as appropriate. This method is particularly usable for small files, or when using such aperture card production systems as the 3M Filmsort Camera Processor.

2. Sequential Filing - Machine-Punched Accession Numbers. The aperture cards are stored in a file, with dividers. Each card is punched with the appropriate identifying information, and a slave deck is created, for subsequent machine processing and searching. Collections of up to 1 million cards are being successfully handled on this basis. Requirements include: (1) the files must not be open to the users, and (2) sequential searching and filing methods should be usable.

3. Random Filing. The aperture cards are stored in a file, with subject-type dividers. Duplicate cards may be used, such that a card appears behind each appropriate subject heading. Or, if desired, only a reference card is filed under alternative subject headings, cross-referencing the location of the aperture card. This approach is very effective for personal files of magazine articles, vendor information, etc.

In the first two cases (which are used by the majority of aperture card users), browsing through the files is not feasible. An external index must be used. Usually, an externally-generated list of cards to be pulled is used, and is ordered in the same sequence as the files.

4.1.2 Semiautomated Systems

Semiautomated systems can be used for aperture cards. These systems have the following characteristics:

1. Files are usually randomly ordered.

2. The aperture cards are modified so as to provide edge-notched codes, rather than internally-punched codes.

3. A keyboard device of some kind is used.

Two systems are typical of this kind of retrieval system:

1. Needlesort....McBee or equivalent.
2. Joggers....Acme Visifile/Remington Power Files, or equivalent.

The essential difference between the two systems is that in one case, all unwanted cards are discarded, and the desired cards are left hanging on the needle; in the other, the wanted cards are shoved out of the pack of unwanted cards.

These systems require that the aperture card be edged with a metal strip, or the film image mounted in a properly-cut card to permit use of the retrieval technique.

Jogging devices are becoming more common, partly because there is less manipulation required of the cards; hence less damage is likely to the film.

Joggers cost on the order of \$1,200 for one tray, holding 750 cards and one keyboard unit. Additional trays can be obtained for about \$600-800 per tray.

Approximately 25 discrete code positions are available on the strip, but only 12 codes are feasible, because 2 code positions are used for each discrete entity.

4.1.3 Fully-Automated Systems

There have been several fully-automated aperture card systems developed, at large costs, that have not proven to be commercially feasible, at least to date. This does not mean the concept of these systems cannot be utilized at some future date. However, consideration will be given only to the Mosler Selectriever System.

The Mosler system is the most widely used large-scale random access, multi-purpose aperture card system. The basic premise is a vertical file of 200,000

aperture cards, each encoded with its own special accession number. When the keyboard is activated, a picker tracks to the proper tray, the card is located, pulled out of the tray and ejected from an output port. The keyboard may be at the file location, or may be located anywhere, including on a touch telephone.

Several output options are possible. For example, closed circuit TV can be employed. The aperture card is positioned where it can be scanned by a television camera, equipped with a zoom lens. At the user's discretion, the entire 35mm frame may be viewed or a detailed section may be scanned.

Other output options are provided, including automatic card duplication, or hard copy generation at the central file location. Or, the aperture card can be ejected, for subsequent manual processing. Regardless of the purpose for which the card is extracted from its stored location, refiling is to the original location of the card. The original location is random, but subsequent filing is to that specific location.

Retrieval times are 3-6 seconds, cost is \$30,000 and up.

Reexamining some of the past systems designs that have not been commercially successful is profitable in that the direction of interest can be clearly established. When the technological developments have reached a sufficient maturation point, the basic approach may prove to be economically sound after all.

Rather than to cite specific systems, it would be better to describe the characteristics¹³ of such a system. That way, we can save time, and still permit consideration of the basic elements.

These systems provide as an output, an aperture card containing the desired image, say an engineering drawing.

The aperture card is not stored, but rather is created from the microform stored originally. This microform may be 35mm roll film, or film chip, 16mm roll film, microfiche, videotape, or any other storage medium, such as binary digits.

When a particular image is required, the search mechanism is initiated. This may involve full search capabilities, on-line browsing, or transference of a reference number. Regardless, the computer-based system searches out the particular image, and creates an aperture card for that stored image. The image may be projected, as for example, onto a cathode ray tube, or onto film, or may be contact-printed onto film. The basic reasons why the systems were unsuccessful ranged from the cost of the hardware and software itself, through the complexities of handling film at high speeds, and included the difficulties of making the aperture card image itself.

Incidentally, some systems today utilize most of the principles of operation we have discussed, except that output from the files is not aperture cards, but roll film or hard copy, etc.

Further, the implications of high-reduction storage, at perhaps, 10,000 full page images per inch, using laser technology suggests that the need for generation of microfilm output from stored images continues to exist.

Some of you may be wondering why machine searching of aperture card files is not more common. Certainly, there's been considerable effort made to punch information into these aperture cards. The fact is that it is not advisable to run aperture cards through machines because the film may become damaged, and because the card, itself, may become damaged. Further, the primary problem is not the selection of the aperture card when the accession number is known, but rather establishing the accession number.

That problem, establishing the accession number, is in fact the primary difficulty in all information retrieval systems.

4.2 MICROFICHE RETRIEVAL SYSTEMS

Microfiche retrieval systems add another dimension to the problems of searching and locating information within the files. In aperture cards, we need find only one thing...a specific aperture card. In microfiche systems, we need to find not only the individual fiche, but sometimes we must also find an individual image within that fiche. We will consider both problems in terms of manual, semiautomated, and automated systems.

4.2.1 Manual Systems

Microfiche retrieval systems embody the same principles of filing for manual retrieval as do aperture card systems. Manual systems usually do not locate individual images, they only locate the fiche. The individual page must be located by putting the fiche into a viewer and searching through the fiche, or by using the viewer coordinate display scheme to match that of the fiche, and to go directly to the preselected image.

Fiche contain a normal-reading accession number (and usually a title block, as well). Dividers are used to separate the fiche into related subsets by the accession number, title, or other classification. These subsets can be searched quickly by visual inspection. External indexes and surrogates are used to locate the accession number. The file is searched:

- a. By subset, marked by file, drawer, divider
- b. By Accession Number

Another frequently-used technique is to place a self-contained index (in reduced size) on the first fiche of each file. Thus, the first fiche of the file, or subset, contains an index to the individual images within that file.

For convenience in filing, and to aid in locating misfiled fiche, duplicate fiche are often color-coded. The fiche is duplicated using the appropriate film in one of many available colors, such as red, for classified information.

4.2.2 Semiautomated Systems

Random access to files of microfiche can be accomplished by the same methods as for aperture cards...use of power files, joggers, and needlesort techniques. Retrieval is of the fiche, and is not to the image level.

4.2.3 Automated Systems

Presently, only one system is widely used for automated search of microfiche files, to the individual image level. That method is the Houston Fearless CARD system (Compact Automatic Retrieval Display). The CARD system is a desk top, self-contained file-reader. Pushbutton selection from the control panel provides 4-second access to any desired image in the file. The file contains up to 750 fiche, and utilizes a 12-bit code, similar to the power file jogger devices mentioned previously. The image is positioned so as to be readable on the screen, or to be copied on hard copy, etc. The unit may be controlled manually, or by computer.

Cost of one unit is about \$3,150.00, without computer control. The fiche may be removed and replaced with other fiche. It may be feasible to hookup several retrieving units to a central display console to increase the capacity per unit.

The Sanders-Diebold SD-500 system was announced in the spring of 1968. The system utilizes Diebold power files, matched with remote viewing terminals, utilizing Sanders display units. The entire system is computer-controlled, and provides access to up to 5,000,000 images at a claimed average retrieval time of 8 seconds. Many different microforms can be mixed within the system, including chips, COSATI and NMA microfiche, film strips or individual frames of 8, 16, 35mm film. Hard copy printout capability for individual images can be provided at the terminal, or at the central power file. Costs vary considerably with the number of terminals, and other potential system differences.

The Mosler Selectriever provides access to 200,000 fiche (or more), when fiche are mounted in the unit in place of aperture cards. Presently, access is to the fiche, and not to the individual frame.

Considerable effort, particularly in Project MAC at MIT is going into computer-retrieval of the individual fiche, and subsequent scanning and browsing of the individual pages on the fiche from removed locations. Much remains to be done in this area to make such a system economically feasible. The biggest problems in this area are transmission of the image through conventional telephone lines (very slow), and multiple access capability, which requires a separate buffer. Buffers are expensive.

4.3 ROLL AND CARTRIDGE SYSTEMS

Cartridge systems are rapidly superseding roll film applications, because cartridges are easier to handle, and protect the film better. Roll film has been mostly 35mm film, cartridges tend to be mostly 16mm film. Newer systems, such as VSMF using 8mm film at a 38x reduction ratio appear very promising.

Cartridges are often color-coded to speed up identification and retrieval. They are marked with an accession number. They are dated, time-ordered, or subject ordered in filing. Most systems employ an external index for the cartridge and frame identification, although some systems do employ an internal index to locate the desired frame. Locating the cartridge is a problem similar to that for aperture cards and microfiche, except that the cartridges must be stored in racks, rather than in files.

The internal indexing techniques range in sophistication from simple frame counters to multi-faceted marks which can be used with a keyboard, providing basic Boolean logic...either/or, and/or, neither, nor, etc. Costs for viewing equipment only ranges from about \$400 to \$40,000, depending on the indexing and retrieval techniques used. Some of the techniques include:

1. Frame counter. The counter is related to the footage, and provides a zone of interest. Fast search speeds are used, until the counter indicates that the desired image is nearly reached. Then individual frame inspection is conducted by the user.

2. Leader Technique. Instead of using a frame counter, a white leader is introduced every 100 to 500 images. The leader can be quickly spotted.
3. Line Marker. This technique employs a black marker on the individual page. The marker is shifted every 100 frames. When the user advances the film, the marker appears as a black line on the viewing screen, and can thus be used to establish the zone of interest.
4. Image Counter. Each frame of the film has a marker, which can be counted by an electronic counter. The image number is entered through a keyboard, and the film is advanced as required to locate the matching number. The desired frame is then displayed on the viewing screen.
5. Blip Coding. This technique employs binary, or binary-coded decimal codes incorporated into the film, usually preceding the image being identified. The equipment permits identification of one or more of the characteristics described in the code to be entered into a keyboard, and subsequently used as a scan comparison, by which the desired image (specifically known or not) is located.

4.4 CHIP SYSTEMS

Several of the larger and more sophisticated retrieval systems use film chips, rather than roll film, aperture card, or fiche. Costs of these systems vary considerably. Costs may range from under \$5,000 to over \$2,000,000. These systems are unique, but contain some features in common:

1. File integrity is preserved. The systems do not permit direct access to individual chips (except when updating/adding to the files).

2. Output tends to be a conventional microform, such as roll film or aperture card.
3. The system incorporates the necessary coding for search logic with the individual chip. Accession numbers, per se, are not used.
4. The system services an organization where the value of the information is extremely high, the amount of information is substantial, and the economic justification for the system is less important than the availability of the information.

4.5 VIDEOfILE

The Ampex Videofile system is not a microfilm system, but rather employs television and computer technologies to provide for document storage and retrieval. The document is stored on magnetic tape in binary language, and has a unique accession number to provide for retrieval. On command, the tape is searched for the desired image, the record is located, and the document is converted from binary language to a display on a television tube. The magnetic recording of the document is erasable.

Capacity of the system is high (unlimited numbers of tapes can be used). Each 7,200 foot tape can store 135,000 standard pages, or 350,000 standard television-sized frames.

Costs vary with the number of tape units, buffers, terminals hooked into one system, but minimum configuration costs would be \$500,000 up to approximately \$2,000,000.

5. SUMMARY

Microfilm information retrieval systems incorporate the principles of information retrieval used in hard copy systems, in manual, or automated systems.

There are two basic types of systems:

1. Those systems in which the microfilm file is addressable, but locating the address is a function of an external process.
2. Those systems in which search capabilities are a function of the microfilm records themselves, in that the images are identified with search information.

The proper system for a particular application can only be determined by looking at the requirements for retrieval, and identifying the coding and storage requirements to permit retrieval. The important thing is to make sure that the user of the information can quickly find what he wants, and that the cost of making that information available to him is less than the cost of not having the information at all.

For further information on information retrieval and microfilm systems, refer to reference numbers 1 through 39 in the bibliography.

BIBLIOGRAPHY

1. Business Automation Magazine, Annual Reference Guide, Volume 15, Number 9, September 1968, 188 pages
2. Guide to Microreproduction Equipment, Hubbard W. Ballou, Third Edition, National Microfilm Association, P.O. Box 386, Annapolis, Maryland, 1965, 552 pages
3. 1966 Supplement to the Guide to Microreproduction Equipment, Hubbard W. Ballou, Third Edition, National Microfilm Association, P.O. Box 386, Annapolis, Maryland, 1966, 127 pages
4. 1967 Supplement to the Guide to Microreproduction Equipment, Hubbard W. Ballou, Third Edition, National Microfilm Association, P.O. Box 386, Annapolis, Maryland, 1967, 115 pages
5. Tutorial Glossary of Documentation Terms, American Documentation Institute, 29th Annual Meeting, October 3-7, 1966, Santa Monica, California, 78 pages
6. A National Library Agency...A Proposal, ALA Bulletin, March 1968, pp. 255-265
7. Survey of IS&R Equipment, Lawrence H. Berul, Datamation, March 1968, pp. 27-32
8. Methods of Information Handling, Charles P. Bourne, John Wiley & Sons, Inc., 1963, 241 pages
9. Principles of Automated Information Retrieval, William F. Williams, The Business Press, Elmhurst, Illinois, 1965, 439 pages
10. Information Retrieval: An Introduction, Robert M. Hayes, Datamation, March 1968, pp. 22-26
11. Management Communication Across Space and Time, Louise Schultz, System Development Corporation, 2500 Colorado Avenue, Santa Monica, California, SP-1860, 15 November 1964, 8 pages
12. Library Fiche: An Introduction and Explanation, Arthur Teplitz, Pre-print-General Session Papers and Seminar Materials for the 17th N.M.A. Convention, May 21-24, 1968, Chicago, Illinois, pp. 117-124
13. A Review of Microphotographic Techniques and Graphic Storage and Retrieval Systems, Alfred S. Tauber, Presented at the Conference on Automation Fundamentals for Special Librarians, Drexel Institute of Technology, November 3, 1967, 11 pages
14. Retrieval of Technical Documents, Lea M. Bohnert, C-E-I-R, Inc., Washington, D.C., CEIR 70 0219 27, AD 664 596, November 1967, 230 pages

15. Information Selection Systems Retrieving Replica Copies: A State-of-the-Art Report, Thomas C. Bagg and Mary Elizabeth Stevens, National Bureau of Standards Technical Note 157, December 31, 1961
16. The Copyright Law As It Relates to National Information Systems and National Programs, Federal Council for Science and Technology, Washington, D.C., July 1967, PB 175 618, 81 pages
17. Re-Evaluation of Microfilm As a Method of Book Storage, Verner W. Clapp and Robert F. Jordan, College and Research Libraries, January 1963, pp. 5-15
18. Facsimile Transmission in Libraries: A State of the Art Survey, Sharon Schatz, Library Resources & Technical Services, Volume 12, Number 1, Winter 1968, pp. 5-15
19. The Bush Rapid Selector, Phil Hirsch, Datamation, June 1965, pp. 56-57
20. Computer Gets Faster Running Mate, Business Week, June 8, 1968, pp. 84-84
21. New Look in Manual Methods, Robert W. Gibson, Jr. and Ben-Ami Lipetz, Special Libraries, Volume 47, Number 3, pp. 106-113
22. The Facts About Facsimile, David H. Azner, Data Processing Magazine, May 1968, pp. 42-53
23. The Storage and Retrieval of "Visuals", Walter S. Baker and Alexander G. Hoshovsky, Graphic Science, March 1968, pp. 22-27
24. Reference It; Retrieve It; Reproduce It, Richard S. Nauer, Systems & Procedures Journal, March-April 1968, pp. 32-36
25. Cost Characteristics of Library Service, Masse Bloomfield, Special Libraries, Volume 58, Number 10, December 1967, pp. 686-691
26. The Validity of Comparative Technical Library Statistics, Margaret N. Sloane, Special Libraries, Volume 58, Number 10, December 1967, pp. 692-696
27. Document vs Digital Storage of Textual Materials for Network Operations, EDUCOM, Bulletin of the Interuniversity Communications Council, Volume 2, Number 6, December 1967, 9 pages
28. An AFIPS Report: The State of the Information Processing Industry, April 1966, 103 pages
29. Information Storage and Retrieval: A State-of-the-Art Report, Lawrence H. Berul, Auerbach Corporation, Philadelphia, Pennsylvania, 14 September 1964, AD 630 089, 200 pages

30. Automatic Indexing: A State-of-the-Art Report, Mary Elizabeth Stevens, 30 March 1965, National Bureau of Standards Monograph 91, 220 pages
31. Information Retrieval, A Critical View, Edited by George Schechter, National Colloquium on Information Retrieval, May 12-13, 1966, Philadelphia, Pennsylvania, Thompson Book Company, Washington, D.C., 1967, 282 pages
32. A Guide to Personal Indexes Using Edge-Notched and Peek-a-Boo Cards, A. C. Foskett, Olive Bingley, London, 1967, 80 pages
33. Proceedings of the American Documentation Institute Annual Meeting, Volume 4, Levels of Interaction Between Man and Information, New York, October 22-27, 1967, 319 pages
34. Proceedings of the Fifteenth Annual Meeting and Convention, Edited by Vernon D. Tate, The National Microfilm Association, Annapolis, Maryland, 1966, 254 pages
35. Annual Review of Information Science and Technology, Volume 2, Carlos A. Cuadra, Editor, John Wiley & Sons, Inc., 1967, 484 pages
36. Proceedings of the American Documentation Institute Annual Meeting, Volume 3, Progress in Information Science and Technology, Santa Monica, California, October 3-7, 1966, 494 pages
37. Systems for the Intellectual Organization of Information Volume VII, Coordinate Indexing, John C. Costello, Jr., Graduate School of Library Science, Rutgers-The State University, New Brunswick, New Jersey, 1966, 218 pages
38. Clue-Word Indexing of Technical Information in a Research Laboratory, Bruce McCreary, IEEE Transactions on Engineering Writing and Speech, Volume EWS 6, Number 1, pp. 28-30
39. How Much Noise in a Role-Free and Link-Free Co-Ordinate Indexing System?, J. F. Blagden, Journal of Documentation, Volume 22, Number 3, September 1966, pp. 203-209

~~Unclassified~~
Security Classification

DOCUMENT CONTROL DATA - R & D		
<i>(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)</i>		
1. ORIGINATING ACTIVITY (Corporate author) System Development Corporation Santa Monica, California		2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP
3. REPORT TITLE Microfilm and Information Retrieval		
4. DESCRIPTIVE NOTES (Type of report and inclusive dates)		
5. AUTHOR(S) (First name, middle initial, last name) Arthur Teplitz		
6. REPORT DATE 29 October 1968	7a. TOTAL NO. OF PAGES 31	7b. NO. OF REFS
8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S) SP-3243	
b. PROJECT NO.	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)	
c.		
d.		
10. DISTRIBUTION STATEMENT Distribution of this document is unlimited		
11. SUPPLEMENTARY NOTES		12. SPONSORING MILITARY ACTIVITY
13. ABSTRACT <p>This paper discusses the principles of information retrieval, considers subject and classification indexing, and describes elements of coding for manual and machine applications. The implications of information retrieval practices on microfilm information retrieval systems are discussed. Characteristics of information retrieval for manual, semiautomated, and automated systems for aperture cards, microfiche, roll film, and chip systems are considered.</p>		

DD FORM 1473
1 NOV 65

~~Unclassified~~
Security Classification

Unclassified

Security Classification

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Information Retrieval Subject indexing Classification indexing Coding Microfilm Microfiche Aperture cards Roll film Chip systems						

Unclassified

Security Classification