OPTICAL DISK WORKSHOP
and
APPLICATIONS OF AUTOMATION IN
ARCHIVES AND RECORDS MANAGEMENT

A report on a trip to the USA in October 1987

Roswitha Rahmy
OPTICAL DISK WORKSHOP
and
APPLICATIONS OF AUTOMATION IN
ARCHIVES AND RECORDS MANAGEMENT
OPTICAL DISK WORKSHOP
and
APPLICATIONS OF AUTOMATION IN
ARCHIVES AND RECORDS MANAGEMENT

A report on a trip to the USA in October 1987

Roswitha Rahmy
CERN Archivist

CERN
TH/SIS
GENEVA 1988
Table of Contents

Schedule .................................................. 1
Introduction .............................................. 1

Part 1: Optical Disk Workshop

Chapter 1: General Remarks. ............................. 2
Chapter 2: Lectures and Systems Demonstrations. ....... 3
  2.1 at NARA. ........................................... 3
    2.1.1 Optical Digital Disk
    2.1.2 Optical Character Recognition
  2.2 at Library of Congress ............................. 5
    2.2.1 Optical Disk
    2.2.2 Video Disk
  2.3 at Filenet ......................................... 6
Chapter 3: Standards .................................... 8

Part 2: ICA Section of Archivists of International Organizations
        SIO - 11th Session

Chapter 1: General Remarks. ............................. 10
Chapter 2: Applications of Automation in Archives and Records
          Management ....................................... 10
  2.1 MARC-AMC Format .................................. 10
  2.2 Applications of MARC-AMC in some U.S. Institutions 11
  2.3 Other Applications of ADP in Archives and Records
          Management ....................................... 12
  2.4 Microcomputers for Archives ....................... 12
  2.5 Foreign Affairs Information System at U.S.
        Department of State ............................. 14
Chapter 3: Summary of SIO-Discussions and Recommendations ... 14

Part 3: Visits to Archives

Chapter 1: Smithsonian Institution, Washington. ........ 15
Chapter 2: American Institute of Physics, New York. .... 16

Conclusion ............................................. 17
Bibliography .......................................... 18
Appendices
Trip to USA from 5 to 15 October 1987

Report by Roswitha Rahmy / TH-SIS

Schedule:

5-7 October: Participation in a Workshop on automation and optical digital disk technology, organized by the International Council on Archives [ICA], Automation Committee at U.S. National Archives and Records Administration [NARA]

8-9 October: Thirteenth session of the ICA, Section of Archivists of International Organizations [SIO], hosted by the International Monetary Fund [IMF]; working paper presented by R. Rahmy

9 October: Joint sessions of the Automation Committee and SIO

12-14 October: Visits (by R. Rahmy) to U.S. National Archives and to Smithsonian Institution Archives

15 October: in New York.
Visit to the American Institute of Physics, Center for History of Physics and Niels Bohr Library.

Introduction

The aim of the trip was to participate actively in the Workshop on optical digital disk technology and the session of the ICA Section of Archivists of International Organizations furthermore to visit and to discuss major archival aspects with archivists of U.S. National Archives and Records Administration, of the Smithsonian Institution and of the American Institute of Physics. It has been possible to establish very good relations with the archivists of these institutions as well as with several colleagues of International Organizations and the CERN archivist will continue to keep in contact with them.

The Optical Disk Workshop was organized in form of theoretical lectures, discussions, systems demonstrations at NARA and visits with demonstrations and discussions to the Library of Congress, the U.S. Department of State and to Filenet, a commercial firm with whom NARA is working for their optical disk project.

In addition to this, lectures were given on standards in archives and records management by representatives of the National Bureau of Standards, Washington and finally a demonstration with discussions was given of the microMARC-AMC system developed on the basis of MARC-AMC by the Michigan State University Archive and implemented there.

The 13th Session of the ICA Section of Archivists of International Organizations was almost entirely devoted to the use of computer systems in archives and records management, to the implications of automation for registry systems, and to optical disk technology. This report is divided into three parts. The first part deals with the Optical Disk Workshop, with systems demonstrations related to the Workshop and with standards applications in archives.
The second part reflects syntheses of SIO discussions on and around the subjects exposed in the working papers and in oral communications on projects in archives and registries of several international organizations. Both parts are preceded by general remarks on the subjects treated later on in more detail.

A small third part is dedicated to visits to two archives after the SIO session and, in a way, completes part 2.

Exhaustive appendices and a bibliography provide more detailed information.

Part 1: Workshop on Optical Digital Disk Technology and Standardization

Chapter 1: General Remarks

The Workshop on Optical Digital Disk Technology and Systems Demonstrations organized by the Automation Committee of the International Council on Archives has been opened to the participants of the ICA Section of Archivists of International Organizations. The Workshop took place at the U.S. National Archives and Records Administration, Washington, from 5-7 October 1987 and was chaired by Charles Dollar, Assistant Director of the NARA Archival Research and Evaluation Staff.

The lectures of the Workshop were supplemented by systems demonstrations and visits to different institutions. Furthermore a lecture was given on standardization of formats by staff of the National Bureau of Standards, Washington.

The advantages of the optical digital disk are:
1) very large storage capacity
2) reliability in repeated use (no abrasion or wear as for magnetic tapes)
3) not damaged by dust, finger prints etc.

It converts images on paper or on film into a computer code. Furthermore the electronic image enhancement at the optimum threshold gives the possibility to scan pale documents and to augment the quality of the copy.

The main reasons to introduce optical disk technology in archives are the following:
- preservation of documents
- access (easier, faster)
- service to users (better)
- space (high storage capacity).

For some applications, cheaper and less labour intensive systems, like Optical Character Recognition (OCR) or Video disks, are used with success.
Standards for data base management, data base languages and data exchange have shown useful, or even necessary in practice. Standards have been developed since the 1960's by ASN and nowadays by ISO.

Chapter 2: Lectures and Systems Demonstrations

2.1 National Archives and Records Administration (NARA) Washington

Acting archivist: Dr. Frank G. Burke

NARA is the repository of the permanently valuable records and documentation of more than 500 agencies of the U.S. Federal Government. The Archival Research and Evaluation Staff Section is pursuing the transformation of NARA holdings to machine processible information databases.

Assistant Director: Charles Dollar
Assistants: William L. Hooton (ODISS)
Mary Allen (OCR)
Thomas Weir (Microcomputers)

2.1.1 Optical Digital Image Storage System (ODISS)
Mr. William Hooten - responsible

ODISS is a test/pilot research project for a selected part of existing records to determine whether a conversion of large paper holdings to an electronic form is feasible for the Archives. Data collected during the pilot project will be published in a detailed report of the test findings, scheduled January 1989.

The assumptions for the ODISS project are the following:

1) the system must accept wide diversity of input of documents and microforms using digital scanning technology
2) long term image storage by using optical digital data disks
3) the system must be able to automatically retrieve the document images stored in an optical library
4) the system must produce "hard copy" output of document images on demand
5) remote access to index data must be possible using appropriate telecommunications capability
6) work-station displays will be user friendly.

The ODISS system is divided into two main operations:

a) electronic input scanning and image storage
b) reference and retrieval.

a) Electronic input subsystem:

- includes the document preparation (and evtl. repair),
- scanning with a high speed automatic scanner for straightforward material and with a low speed manual scanner for delicate handling,
- assignment of a unique file control number,
- indexing,
- quality control (100%).
b) Reference/retrieval subsystem:
three areas of reference activities are foreseen

- search by NARA staff
- public searches (in-house)
- off-site index searches with normal (i.e. not a high resolution
  graphic) terminal and a printer at Tennessee State Archives, Nashville.

ODISS, implemented at NARA on a trial phase, responds to the following
criteria:

- retrieval is easy, rapid
- increased document legibility
- indirect preservation of documents
- faithful reproductions.

Paper and microfilm can be scanned. NARA is using a TDC-DS-4000 scanner
produced by Terminal Data Corporation, Woodland Hills, CA). This scanner
is compatible with any computer.
The storage is done on a Sony optical disk with high density scanning
(200 lines vertical is common average) for high technology projects.
Indexing is done by special UNIX-based software tailored for NARA.
Quality and indexing control is first done on tape (buffers) and after
control and evtl. rescanning of faulty images the tape is put on optical
disk by Sony in Japan.

To copy a digital optical disk onto another one takes about 2 hours.
Mylar plastic folders are used for fragile documents when scanned.

2.1.2 Optical Character Recognition (OCR)
Mrs. Mary Allen - responsible

OCR implies automated conversion of written data into machine-readable
form.
NARA uses this pilot project for "low", which means inexpensive, technology
projects on a research test contract.
It works on IBM-PC with a DEST Desktop-reader and PC-scanner (quite an
inexpensive machine).
DEST-reader reads 12 kinds of typefaces, transforms into ASCII, but does
not retain images. It reads at a speed of 15-20 sec./page

![Desktop-reader OCR diagram]

Search with Boolean operators, possibility of full-text retrieval.
System can create "Folders" (=subfiles) for subject research.
Software: Textbank on IBM-PC

Textbank is an off-the-shelf commercial software available from:

Groupe L Corporation
481 Car lisle Drive
Herndon, VA 22070

(500.- $ for use on IBM-PC; possibility to upload to a big IBM, software might be more expensive).

Storage is on magnetic tape.

Comparison

Optical Disk
"dumb copying"

High tech. storage tool for image scanning, copies lines like a photograph, no search on disk without additional manual indexing

Optical Character Recognition
"smart copying"

Low tech. storage tool, reads words, therefore possible to do full text retrieval without additional indexing

2.2 Library of Congress, Washington

The Library of Congress holds a collection of about 80 million items with a growth rate of 7,000 new items per day.

In 1982 the LC began research in the use of optical disk technology for image preservation and retrieval and later on also for use in the Cataloguing Distribution Service. The LC Optical Disk Storage Technology Committee has chosen two experimental systems:

1. Optical digital disk system from Integrated Automation

2. Video disk and compact audio disk system from Sony Video Communication Products

Director of Planning Office: Bob Cich

2.2.1 LC Optical Disk System

Pilot project since 1982, final report expected in Oct. 1988; to be published later. Also test results will be published.

They use page scanners and microfiche and microfilm scanners. Disks are gold alloy 12" disks of 1 Gigabyte. Scanning resolution is 300 lines/inch.

Before scanning the material the copyright question must be clarified. LC is Copyright Depository Library.
Material put into the system:
- Articles from periodicals or complete journals
- Manuscript file
- Maps
- Sheet music
- Congressional reports

Cost is about twice as high microfilm technology but improvement in service might be worth while to have higher cost.

It is a specially conceived system for LC by Integrated Automation System, Alameda, California and uses laser disk technology.
A one-sided 12" disk can store 15000 images of double space typing but only ~4000 images with dense print.
A jukebox (holding 120 disks) and buffer memories are used. Multiple access is possible (many terminals scattered all over the building). 15000 pages per day can be scanned.
The indexer also makes quality control.

As this is a pilot system, no back-up disks are made.

2.2.2 Video Disk System

LC uses this system with low resolution for motion image disks and photoalbums. Colour photos of posters are taken on 35 mm film and than transferred to videodisks (54000 images/disk) by Sony in Japan.

Cataloguing: of captions by BRS.
Keywords of LC-thesaurus are used for graphical material.
Every photograph or poster gets an identification number and then gets catalogued apart.

Search: Sequential search on photographs via microcomputers using a UNIX-based retrieval system.

1. The system is menu-driven
2. Records contain the information given in the old catalogue.

Print: one can make quick and inexpensive black and white prints near to the search station.

2.3 FileNet

in Bethesda, Washington.

File Net is the name of a commercial firm which sells an optical disk-based storage and retrieval system for active records management. In Europe FileNet is represented by Olivetti (OliNet). It is an integrated hardware/software system which converts paper documents into digitized images stored on optical disks. The software is delivered with the system.
NARA works with the FileNet system on their pilot project ODISS. There is also the possibility to use already existing information on another computer mainframe (i.e. on IBM).
The FileNet system consists of four basic parts which are interconnected by an Ethernet Local Area Network (LAN):
- document entry
- image management system
- document retrieval
- printing

2.3.1 Document entry process:
the documents are fed into a scanner with the resolution of 200 or 400 dots/per inch. A document entry server drives the software for inspecting and indexing the document-images. The image is visualized on a terminal, with a screen of 22 inches wide and 115 dots/per inch, for quality control and short (keyed) indexing.

2.3.2 Image management system:
the indexed images are stored on non erasable optical disks of 12 inches, with an information layer of tellurium alloy between 2 glass layers, and 10 years guaranteed lifetime. Storage capacity is 52000 pages A4 (or 2.6 gigabytes) on each dual-sided disk.
On-line storage: in an Optical Cabinet (52000 pages A4)
Otherwise storage is done in:
- Jukebox OSAR=64 (64 x 52000 pages)
- Jukebox OSAR=200 (200 x 52000 pages)

There is, of course, a possibility of additional off-line storage.

The system uses a specialized software analogous to a data base management system(DBMS). It provides support for electronic "file folder" retrieval and password security.

2.3.3 Document retrieval

The processing and retrievals done at one or several integrated work-stations with specific software for windowing, text edit, retrieval with Boolean logic and optional access to another computer mainframe (IBM, VAX). There exists also an optional software package called WorkFlo which accelerates and simplifies all document-image processing. The split-screen gives the possibility to show several images (up to 16) side by side.

2.3.4 Printing

Printing is done on a laser printer at the rate of 12 pages/per minute. It is high quality with a printing density of 400 dots/per inch.

2.3.5 The system produces automatically on a spare optical disk a security copy of all text and images. It allows simultaneous multiple user access.

Although no direct figures on cost could be obtained, it seems that the system is very labour intensive and expensive. It is working in big insurance, banking and investment companies.
Chapter 3: Standards

3.1 General Remarks

The problem which the archives have to face is the development of a policy for using existing and emerging international standards as tools for the transfer of electronic databases from the information-creating agencies to the archives and for the access to these databases for the users.

The data of archival value must be available to the archives in a format which facilitates the transfer to the archive and the storage for future retrieval. As technology of automated information systems changes quite rapidly, the need for standardization has become more and more obvious.

The International Organization for Standardizations-ISO - is the major standards writing body, but only a small part of its efforts is applied to data communication. ISO is divided into Technical Committees (TC) that write standards in particular fields:

TC 46 - Documentation
TC 97 - Information Processing
TC 154 - Documents and Data Elements in Administration, Commerce, and Industry
TC 184 - Industrial Automation Systems.

For archives TC 46 and TC 97 are of greatest interest. Objectives for standardization in archives are:
- requirements for transfer of documentation
  (specify interchange form and contents)
- prespecification of transfer structures (formats)
- reformatting of archival applications
  (creation of finding aids and keyword lists etc.)
- high degree of access provided to users
  (user friendly software, easy retrieval etc.)

3.2 Lecture- National Bureau of Standards

As a complementary source of information to the participants of the Workshop, lectures on the use of standards for the preservation of electronic records have been delivered by Leonard Gallagher and Larry Welsh of the National Bureau of Standards, Washington, under the aspects of database considerations and document considerations.

3.2.1 Database Considerations (L. Gallagher)

Historically Random Access Data models in the 1950's preceded standards for Network Databases which started in the 1960's, followed by Relational Databases in the 1970's and by Relational Interfaces in the 1980's.

The ISO adoption of DBMS standards (Data Base Management System) dates from the latter part of the 1980's.
The following standards for Database Management exist:

- Database Language NDL (Network model):

- Database Language SQL (Relational Model):

- Information Resource Dictionary (IRDS):
  ISO DP 9075 (ANSI X3.138-1988)

Standards for Data Interchange are:

- Data Descriptive File (DDF):
  ANSI/ISO 8211 - 1985

- Abstract Syntax Notation One (ASN.1):
  ISO 8824 - 1988 (ASN.1) to be
  ISO 8825 - 1988 (Encoding Rules) issued

There are possibilities of Database Transfer but standards do not yet exist.

3.2.2 Document Considerations (L. Welsch)

Historic development:
before 1964 - Typewriters
1964-1975 - Document formatters (as Runoff, Format, Pub, Scribe)
1975-1980 - Typesetters (Troff, TEX)

Problems in information exchange have started since 1980 as simple exchanges are no longer possible, the layout of documents is more complicated and standards are needed.

Existing standards are:

SGML - Standard Mark-up Definition Language
       (a logical content document description language)

ODA - Office Document Architecture
       (logical content description and layout description)

ODIF - Format for exchanging ODA documents.
Chapter 1: General Remarks

The Thirteenth Session of the ICA Section of Archivists of International Organizations took place on 7-9 October at the International Monetary Fund in Washington and was devoted to automation in archives and records management.

As the use of microcomputers in archives becomes more and more common, particularly in relation with word processing and ad hoc applications in smaller archival units, a lecture on microcomputers and their application at NARA completed discussions on this subject.

Furthermore a visit to U.S. Department of State was organized for a demonstration of the Foreign Affairs Information System.

Several working papers have been presented and discussed, in particular on MARC-AMC, on the computerized retrieval and management system of the CERN Archive, and on the automation of the Registry of WHO Geneva (see appendixes 1, 2 and 3). Some recommendations to the International Council on Archives have been adopted (see appendix 5).

One major topic discussed at this Session was the MARC-AMC format, which is the only existing basis for standardization of electronic records in archives and records management. Therefore this report will first expose more general aspects of MARC-AMC, whereas the more detailed description is given in appendix 1.

Chapter 2: Applications of Automation in Archives and Records Management

2.1 MARC-AMC

2.1.1 Generalities

MARC is an acronym for Machine Readable Cataloguing and was established by Library of Congress in 1982. It is a format for information exchange and has been adopted in the United States and in Canada in 1982. In 1983, a joint effort of the Library of Congress and the Society of American Archivists National Information Systems Task Force produced the revised and expanded MARC-AMC (Archival and Manuscripts Control) format.

2.1.2 As a subset of the American National Standards Institute (ANSI) and of the International Standards Organization (ISO) MARC is an internationally accepted information exchange format. The MARC-AMC is a professional standard for the description and exchange of archival information, including both traditional descriptive fields for the contents of records and newer process control fields for noting actions taken on records.

2.1.3 While the format provides a technical structure for exchanging data and a framework for organizing data about archival material, the determination how to use particular data elements in the format (or whether to use specific ones at all) is left to each institution according to its own requirements and individual interpretation of the format.
2.1.4 The MARC-AMC format is independent of hardware and software and allows the exchange of automated information among computer systems in a standardized way. It consists of seventy seven data fields of variable length and each data field has 1 to 20 subfields. [The description of the fields is given in appendix 2, which is Working paper SIO/87/WP-3 presented by A. Erlandsson at the 13th session of the ICA Section of Archivists of International Organizations, Washington D.C., 7-9 October 1987, 15 pages.]

2.2 Applications of MARC-AMC in some US institutions

2.2.1 The Research Libraries Group (RLG) runs the Research Libraries Information Network (RLIN), a data base and resource sharing tool that incorporates the MARC-AMC format.

2.2.2 The American Institute of Physics, Center for History of Physics in Washington is working on a project to automate its "International Catalogue of Sources for History of Physics and Allied Sciences" so that records may be shared with RLIN-AMC (see paper by Joan Warnow, The Reconversion of the International Catalogue of Sources for History of Physics and Allied Sciences, Proposal (AIP Center for History of Physics, October 1987. 29 p. + appendix).

2.2.3 In Spring 1984 the Archival and Evaluation Staff of the US National Archives and Records Administration (NARA) conducted a study of "The MARC Format and Life Cycle Tracking at the National Archives", published in 1986, with the aim to determine:

- the implications of developing an automated system for controlling Federal records throughout their life cycle
- the full capability of MARC-AMC to carry information for control of and access to Federal records
- the usefulness of RLIN-AMC to meet the NARA requirements.

The findings of the study were that:

- NARA should adopt the field definition of MARC-AMC as far as possible to form the basis of a local systems design and development which would enable NARA to enhance its own standards and retain compatibility with external data bases. As predictable, MARC-AMC does not meet every need of NARA (for instance, insufficient hierarchical structure).

- Life cycle tracking requires exchange of information between offices inside NARA and MARC-AMC is a standard mean for exchanging data elements.

- RLIN-AMC is a major bibliographic resource for NARA archival information. Placing information about NARA holdings into RLIN is equivalent to publication, with much greater efficiency and reduced time and expense.
2.2.4 MicroMARC:amc

MicroMARC:amc is a microcomputer system based on the US MARC-AMC format. It has been developed in the Archives of the Michigan State University at East Lansing and is now implemented in 17 other U.S. archives (see appendix 4).

A demonstration programme and Tutorial was presented by Mr. Frederick L. Honhart, Director of the University Archives and Historical Collections of the Michigan State University. It is a local stand-alone system using MS-DOS, which can be up-loaded to a mainframe using the U.S. MARC-format. The system is sold by the Archives and a demonstration diskette with a documentation can be purchased. The system seems to be very flexible but quite slow and bulky for input and search.

2.3 Other Applications of ADP in Archives and Records Management

2.3.1 The implications of automation for Registry Systems have been exposed by Lisa Fagerlund, WHO GENEVA, in her working paper S10/87/WP-1 (Appendix 4). L. Fagerlund showed that the introduction of electronic records management very much simplified the complex handling of documents and their indexing and allowed to come back to an identification and control at document level. A CARDBOX software is working on IBM-PC.

2.3.2 Working paper S10/87/WP-8 on the computerized archive management and retrieval system presented by Roswitha Rahmy, CERN Geneva, explained a fully working system using INFOIL-2 software on an IBM-mainframe computer. (Appendix 3).

2.3.3 At NATO, Bruxelles a pilot study is going on for the next 6 years on computer-aided translations. The translation is done with an automated dictionary using a Canadian software package. They also foresee a fulltext Optical Disk System in the future (Franz Egger, Central Registry).

2.3.4 The automated system in the Records and Word Processing Center of INTELSAT Washington is working on a big IBM computer, but software is also available for VAX and WANG (Friscilla Ruddiman).

2.4 Microcomputers for Archives (Thomas E. Weir, NARA)

Ted Weir gave an introductory lecture on the use of microcomputers in archives and a demonstration of the system used at NARA.

2.4.1 The planning which is necessary for the application of a microcomputer in archives is related to the following points:

Strategic planning needs:
 a) Institutional analysis
 b) Institutional goal setting

Systems analysis:
 a) Study current systems
 b) Identify possible approaches
 c) Select preferred solutions
 d) Install system, resolve problems, begin operation
 e) Maintain system (including reviews)
Basics:
  a) Central processing unit
  b) Random access memory
  c) Input device
  d) Monitor
  e) Mass storage device (hard disk)
  f) Operating System
  g) Applications software

2.4.2 The most common microcomputer operating systems are:

  a) MS-DOS (IBM-PC compatible)
  b) OS II
  c) Macintosh finder
  d) UNIX (multi-users system, needs professional programmer's support)

2.4.3 Software types

Word processing (with specific software):
  a) Word Perfect
  b) Display Write 4 (IBM)
  c) Multimate (Wang)
  d) Microsoft Word (IBM-PC)

Data Base Managers:
  a) File managers
  b) Form oriented relational data bases
  c) Programming oriented data bases
  d) Large data bases (Unify/UNIX, CDS/ISI etc.)

Management Software

Miscellaneous Software:
  File handler
  Backup
  Front end    Utility programmes
  Macros
  Sidekick

2.4.4 Microcomputers at NARA

Which microcomputer system is used at NARA?

Hardware: PRIME (multiple computers system)
Software: INFO (using COBOL)

Microcomputers are used at NARA since about 2 years but mainly for
wordprocessing purposes. There are ad hoc applications in each section
(House of Senat etc.)
Some computerization exists for batch descriptions of series. Detailed
descriptions of folders might be done later on by Optical Character
Recognition.
There are no other microcomputer applications at NARA, yet.
Standards used:
MARC-AMC format is only used as an exchange format:
NARA designed their system for description of series near the MARC-format
and then transform the parts to be exchanged into MARC-AMC.
No ISO-standards are applied yet.

2.5 Foreign Affairs Information Management Center
of U.S. Department of State

Director: William H. Price

The Foreign Affairs Information System (FAIS) is working in a WANG
environment using a commercial system ORBIT as a basis. It was implemented
in 1973 as an automated follow-on to the paper-based central files of the
Department. It captures and maintains information in official
correspondence and communications transacting the business of the
Department. These sources include telegrams, reports from Missions,
diplomatic notes, treaties, Congressional and other correspondence.

The contents of these are indexed by subjects, events, issues and
personalities to provide automated retrieval and access to this
information. Additionally they can be accessed by originator, addresses,
date and security classification.

Full text of telegrams and electronic based correspondence is kept
on-line for direct access for 3 years. Subsequently, this information is
available from magnetic tapes or microfilm. Non-electronic correspondence
is captured on microfilm. Full citation references and index terms are
maintained in the automated system to provide the same search and retrieval
facilities for both electronic and film based information.

Technical support of the system is ORBIT using an IBM 3081 mainframe
computer.

A lot of future improvements in information support is foreseen.

Chapter 3: Summary on SIO Discussions and Recommendations

The working papers presented at the 11th Session were discussed.

In particular with reference to MARC-AMC, members of SIO felt that it
might be too early to recommend the adoption of this format for all
archives of international organizations.

Concerning optical digital disk technologies the idea was put forward
that several smaller pilot projects should be launched in various archives
to gather more information and practical experience as to evaluate the
efficiency and cost for implementation in archives and records
managements.

SIO members were informed that the Advisory Committee for the
Co-ordination of Information Systems (ACCIS), c/o UN Geneva, has proposed
the creation of a Technical Panel on Records Management chaired by
Mr. Richard Barry, World Bank, Washington. This panel would have to study
the impact of the new information and communication technologies on the
especially paper-supported environment of archives and records management
in the organizations of the UN-family (appendix 6).
Awareness and concern that the new electronic information handling technologies will have a wide influence on the traditional activities of archives and records management have determined the Section of Archivists of International Organizations—SI0—and UNESCO as a sponsor, to mandate Mr. Charles Dollar with an extensive "Study on Electronic Records Management and Archives in International Organizations: a RAMP study with Guidelines. Paris, UNESCO, 1986. (PGI-86/WS/12)".

In this study it is clearly noted that archivists and records managers have at present very little influence in the definition of information technology policy of their institutions. They should become involved in decision making on information technology policy right from the sources, as the handling of electronic information will become more and more part of their activities.

Furthermore in an Integrated Office Automation System (including an electronic mail system) an organization-wide policy of principles and rules which govern the management of all recorded information is necessary.

In view of these enlarged fields of future activities for archivists and records managers and in view of the application of automation and inherent new digital technologies SI0 has adopted a series of recommendations to the Executive Board of the International Council on Archives (appendix 5).

Part 3: Visits to Archives

Chapter 1: Smithsonian Institution Archives, Washington
Dr. William Moss, Archivist
Alan L. Bain, Assistant Archivist

The Smithsonian Institution is partly a governmental, partly a private organization and comprises about 15 different museums, galleries and institutes. Whereas parts of the Institution exist already since the middle of the 19th century, the Smithsonian Institution Archives were created in its current form only in 1967, as each organization of the Institution was and still is working in a relatively independent way.

The Smithsonian Archives are a repository of papers of historic value about the Smithsonian Institution and the fields of science, art, history, and the humanities.

Computer applications in the Smithsonian Archives:
only used in archival management
IBM-PC's: for different word processing systems.
Also used in Finance section for budgets (external to Archives).
A history project uses SAMMA for transcripts of oral tapes.
They start to index from manually established finding aids.
The oral history (interviews): all of them get transcripts and a general range of topics covered by interviews is established but they are not really indexed.
There is a project for a photographic collection to be put on a MARC-VM pilot system.
Smithsonian Archives are somewhat committed to use MARC-AMC, but they are not convinced of the format to be satisfactory. In MARC-AMC the fields tend to be insufficient for archival purposes (difficulties in the hierarchical links of the fields). MARC was conceived for libraries, a
fact which still appears in the AMC format. The nearer the archival fields are to books the better one can use MARC-AMC.

Mr. Bain will send RR documentation on Records cataloguing and a Guide. RR should write to Mr. Bain by May/June 88 to get a report on the Systems Study done by a Consultant firm for the Smithsonian Archives (concerning which computer system and which software would suit best the needs of Smithsonian Institution).

Chapter 2: American Institute of Physics, New York
Mr. Spencer Weart, Management Director
Mrs. Joan Warnow, Assistant Management Director

I visited the Center for History of Science and the Niels Bohr Library in AIP. Miss Jean Hrichus, Librarian/Archivist.

The Niels Bohr Library is part of the Center for History of Physics and general archival practices in the handling of the documents in the library are applied. It is a small library, quite unknown to the public, using computers for wordprocessing of transcripts of interviews. No search on computer can be done, yet.

In the Center for History of Science a wordprocessor DEC Decmate III is used (DEC=Digital Equipment Corporation). They are starting automation (about 1000 files) using a microcomputer NEC and as software: RBase.

The system has been tailored for AIP-Center for History of Science by S. Weart. It works with Menu and Tables (somewhat similar to ORACLE). Indexing for the Library and the Archive works as far as possible on standards elaborated by Library of Congress. They are supplemented by Physics Abstracts Classification (up to 1970's). In 1980's there have been AIP classification standards which have become an ISO-standard (J. Warnow dixit). No keywords, no thesaurus exist, yet. The AIP system might become a pilot project.

MARC-AMC format will be used to link to RLIN (Research Libraries Information Network) which include also books information and allows a link to Research Libraries Group.

They also use a Canon PC Printer 80 in the Library for hard copies from 16mm or 35mm microfilms and from microfiches. It has a zoom, the table is orientable, reads from reels and cartridges. It uses normal typing paper and a cartridge of chemicals which produces dry copies. Quite expensive: about 4000.- $.

Joan Warnow is interested in exchange of information with European institutes. She will send proposals. She also suggested to ICA to create a Group of Archives in Science and Technology acting as a platform for information (for the ICA meeting in Paris in 1988).
Conclusion

Archivists and records managers of International Organizations have become aware of the importance to be involved in the automation planning of their institutions. They have also noted the absence of guidelines for electronic information, such as the adoption of the life cycle concept of information and early appraisal of electronic records.\(^1\)

The introduction of micro- and minicomputers, the development of integrated information systems, the growth in Artificial Intelligence applications and optical information storage will increase the complexity and scope of responsibilities archivists and records managers will have to face in the future.

In particular, the increased use of microcomputers leads to a greater decentralization of record keeping and introduces new aspects of the problems of privacy and security.

To face at least part of the problems to be solved, an increased trend to international standardization can be noted. The advantages of using standards are multiple in relation with lower costs in operations, maintenance and training. U.S. MARC-AMC (and the variant MicroMARC:amc) is the only standard which exists nowadays for the formatting of archive files and archive management. It seems to be sufficiently flexible so that the format can be adapted to the needs of the institution. MARC-AMC should be implemented to permit the easy transfer of information as well as some of the ISO standards should be applied.

The increased use of optical information storage will augment storage capacity and speed, as well for input as for retrieval. This is particularly the case for the Optical Digital Disk although only pilot projects are running and no cost/efficiency evaluation exists, yet. The system is labour intensive and seems to be expensive. It therefore needs a careful analysis of systems requirements and a very good feasibility study of the real needs of the institution before introducing an Optical Disk System. The Optical Character Recognition system is less expensive but has not the same applications. It has no high resolution and is therefore not adequate for conservation purposes of documents whereas it is a comprehensive system which interprets black pattern and converts it into machine-readable form.

A lot of projects of application of automation in archives and records managements of International Organizations are on their way. The management of SIO and the ICA Executive Board will function as a kind of "Clearing house" of ideas and of information on studies on automation applications and cost/efficiency evaluations.

\(^1\) See also: Dollar, Charles M. Electronic Records Management and Archives in International Organizations:...
Bibliography

In addition to the lecture notes distributed at the Workshop and the systems demonstrations I have consulted the following items:


3. - INFOL-2; Manuel de référence. Centre Universitaire d'Informatique, Université de Genève et CERN, Genève. 3e édition, 1979.


THIRTEEN SESSION
(IMF Headquarters, Wash. DC, Oct. 7-9 1987)

WORKING PAPER ON ITEM 1 OF THE PROVISIONAL AGENDA
(MARC-AMC - FORMAT)

Introduction

The use of a format such as the MARC-AMC is mainly a question of standardization. The importance of standardization of automation applications in Archives and Records Management has most recently been underscored by Charles Dollar in his RAMP study "Electronic Records Management and Archives in International Organizations" (1986).

Incompatible word processing systems make transmission of documents from one wordprocessor to another difficult. Huge databases cannot communicate with a network of institutional users because of a lack of standardized format and procedures. These consequences can be avoided through the creation of files having a standardized format and which are both media and system independent. This is particularly necessary during database conversion from one computer system to another and where the two computer-systems are not compatible. One such standardized format, specially developed for the archival and records management environment, is the MARC-AMC format.

\[ \text{(1) FORMAT (in general):} \]

The arrangement and coding of information in a machine-readable record in the computer's output.

The MARC-AMC FORMAT: Rules for the standard of information concerning archives and contained in archives. These rules constitute a FORMAT and establish which kinds of information are acceptable and how they should be described. The MARC-AMC Format complies with international standards for information exchange and is independent of any particular hardware or software. It can describe records at any level (group, series and item levels) in terms of content and actions taken for their control.

\[ .../ \]
Because the MARC-AMC format is independent of hardware and software, it allows the transfer of archival databases and files from one operating system to another, as well as the exchange of automated information between different institutions and users.

Exchange-needs in International Organizations

Exchange of automated (electronic) information among computer systems is potentially needed in two situations common to all International Organizations.

The first involves exchange of data among different archival institutions. The reasons are several. Exchanges may be necessary for administrative purposes, or in order to identify what portion of each institution's holdings are open for public research. Information sharing among closely-related organizations within the same system can facilitate the determination of the degree of data redundancy, or the extent to which the archival holdings of one institution duplicate those of another, for the purpose of appraisal. The potential for such exchanges between research institutions and the archives of International Organizations is obvious.

A second application for standardized format exchange is within the context of records management. Many International Organizations have replaced their central registries, which constituted huge manual databases, with a decentralized network of file stations capable of servicing an office or department. When these file stations are automated, a standardized format should be used so that information will not only be exchanged among them, but so that the electronic records thereby produced can be transferred to a central control point. If this is done, automatic life-cycle management can be applied to all records, regardless of form, and be maintained in the manner we will observe through a demonstration of the "Electronic Envelope" concept now used at the United States State Department¹. Whether or not the MARC-AMC format is the best medium for this application is subject to discussion; however, from a long term perspective, this particular format would be most useful when these records become part of an archival holding. If adopted by an International Organization, a life-cycle tracking system such as that allowed by MARC-AMC format could achieve the following goals within the Records Management and Archives functions:

a) create an "audit trail" for a particular records series, regardless of location or office of control, by drawing together information on all actions applied to it over a period of time;

b) keep a complete central inventory of all files within a decentralized system, providing all potential users with references to requested information;

¹ A part of the programme of 13th Session
c) control inter-office loans of paper files,

d) make "horizontal searches" over different file stations,

e) indicate locations of material documenting the same function,

f) indicate when records are scheduled for transfer,

g) provide disposition authority,

h) identify restrictions etc..

Applications of the MARC-AMC Format

To date, the most ambitious and publicized endeavour involving the
MARC-AMC format in an archival/records management context is the Life-Cycle
Tracking project at the United States National Archives and Records
Administration (NARA), aimed at developing an automated system for controlling
Federal Records throughout their life-cycle. Its project managers agreed that
all data elements\(^1\) gathered from various NARA databases had to be
compatible. Such compatibility requires that information used by more than
one office fit into well defined fields, share uniform data element
definitions and be transmitted through a standardized exchange format. These
requirements led to the choice of the MARC-AMC format for NARA's Life-Cycle
Tracking project. The study concluded that once the Records
Management/Archives office had devised a basis for determining the information
and records control needs of offices in an organization, it could then prepare
a set of functional requirements for an automated records system that would
meet the needs of current records, as well as archival, management. System
requirements of this nature will greatly benefit an Organization and place the
Archives and Records Management function in a position of providing a much
desired service to its "constituency", acting as a standard-setting mechanism
for information generated by an Organization and assuring the portability of
that information for transfer to its archives' database. An automated
life-cycle tracking system would place the Archives/Records Management office
in an unparalleled position to provide information about the location and
content of records kept throughout an Organization.

What is the MARC-AMC Format?

At its most elementary level the format is a "container" for
information - a series of labeled pigeon holes - into which data or
information about records may be placed. Within this general framework, users
have the freedom to include various kinds of datafields necessary for their
particular situation and do not have to deal with the mandatory parameters for
computer processing of the records (such as "the leader" or the "records

---

\(^{1}\) Example: Records title, description, creating office, location,
volume, restrictions, transfer date, archives group, administrative
history, series title etc..
The heart of the format is seventy seven (77) variable data fields (see att. 1-2). Each data field has 1-20 subfields, providing considerable latitude in decisions on how subfield information is to appear.

The adoption of the MARC-AMC format may mandate the conversion of many traditional methods of archival description into new and more disciplined and standardized procedures. Successful implementation will depend on the willingness of archivists/records managers to adopt standardized methods and procedures and on the availability of computer programmes to manipulate formatted data.

Alf M.E. Erlandsson
United Nations Archives
MARC-AMC Variable Length Fields (From: "The MARC Format and Life-Cycle Tracking at the (U.S.) National Archives". (1985)

From an archivist's perspective, the Heart of the MARC-AMC record is the variable length fields. These fields generally carry information in narrative form about the material described.

Each variable length field is identified by a three-digit tag (e.g., 245) and by a name (e.g., 245 = Title).

The variable length fields fall into several broad categories.1/

- Fields 0XX contain coded or numerical information, such as system control numbers.
- Fields 1XX contain the name of the person or corporate body that created the material being described.
- Fields 2XX contain the title of the material being described.
- Fields 3XX contain information about the physical attributes of the materials being described, such as a linear measurement.
- Fields 5XX contain a wide variety of narrative notes about the material being described, such as a scope note or a restrictions note.
- Fields 6XX contain subject terms.
- Fields 7XX contain information about additional persons or organizations that created the material being described.
- Fields 8XX contain a variety of disparate information, such as the address of the repository.
- Fields 9XX (if they are used at all) are locally defined.

Each variable length field may be (and usually is) divided into subfields. Each subfield generally is of variable length and carries a part of the information held by the field. For example, the field indicating that a microform is available includes subfields for the type of microform available (e.g., microfilm or microfiche), the source, the availability condition (e.g., for sale), and the order number. As is the case with fields, most subfields within a field are not mandatory.

1/ In this report a number with the form 5XX, 6XX, etc., means any MARC-AMC field between 500 and 599, 600 and 699, etc. Not all numbers (e.g., 689) are used.
Definitions of Major Variable Length Fields

MARC-AMC defines a wide variety of fields for potential use by archivists.1/ To understand the uses of an AMC record, some familiarity with the various fields is necessary.

The variety of fields exist for several reasons. Numerous and well-defined fields encourage standardization of description so that users of the system will easily recognize certain types of information. Breaking data into fields allows indexing to be more accurate and more efficient by ensuring that indexes include only needed information and by helping to define that information. Furthermore, dividing data among fields that can be displayed or suppressed on a screen, as appropriate, allows for the precise and meaningful retrieval of only the desired information.

Each field has a tag number, which is combined with a name to identify the field. In MARC-AMC, the tag number indicates the type of data included in the field. The list below gives the tag numbers for the most frequently used fields and brief definitions of the fields.

Control Fields (0XX)

Control fields contain identification numbers to control AMC records.

001 System control number
A unique identifying number for each MARC-AMC record.

035 Local system control number
Any local control number entered into this field. Use would depend on local systems design.

09X Local call numbers
Ten fields allowing for the storage and searching of local retrieval information. Examples of such information might include appraisal job numbers, agency manual citations, records center accession numbers, and NARS-Al control strings.

1/ A full list of fields can be found in attachment 2.
Main Entries (1XX)

A main entry is the "author" or creator of materials. The concept includes both persons and corporations. Historically, the term derives from card catalogs where one card was the "main entry", other cards were "added entries". In an on-line system, only one "card"exists, but it can be found in many ways. The term "main entry" is, therefore, a little antiquated. The formation of the names for main entry fields is governed by the Anglo-American Cataloging Rules, 2nd ed. (AACR2).

100 Main entry - personal name

Contains the name of the person who created the records (in the case of manuscript collections, i.e., the donor's name).

110 Main entry - corporate name

Contains the name of the corporate body that created the records. The definition of a corporate body is fairly wide and includes such things as ships and convents as well as business corporations and government agencies.

Titles (2XX)

The titles fields allow for a variety of titles. Field 245 contains the principle title of the material being described.

242 Translation of title by cataloging agency

Gives the translation if a foreign-language title is used in field 245.

245 Title Statement

Contains the title of the records described. The title may be the one used by the agency or the archives or both.

Physical Description (3XX)

These fields contain information about the physical attributes of the material being described.
300 Physical description
Contains information about the quantity of material. Measurements may be linear or cubic as needed. Subfields are included for item counts and identification of the item types.

340 Medium
Contains subfields used to identify particular media, which generally are nontextual. The subfields are especially appropriate for machine-readable materials. (In addition to this field in MARC-AMC, there are separate MARC formats for nontextual records). The subfields cover material base and configuration, dimensions, material applied to surface, information recording technique; support; production rate/ratio; location within medium; technical specifications; and material specified.

351 Organization and arrangement
Contains the arrangement statement for a series or a description of the organization of a larger body of material.

Note Fields (5XX)
The note fields generally contain narrative notes that describe the material. Although the note fields have definitions, the actual data entered is at the discretion of the archivist entering the information.

500 General note
Contains information not otherwise accounted for in the other note fields. (See also 59X note fields below).

506 Restrictions on access
Indicates whether there are restrictions on access to materials. This field gives regulatory and legal restrictions such as security classification or physical restrictions (e.g., off-side storage). (For restrictions on use, see field 540 below).

520 Summary, abstract, annotation, scope etc., note
Contains a narrative description of the records. (See field 545 for a narrative description of an agency).
524 Preferred citation of described materials

Instructions for researchers on how to cite materials.

Reproduction Availability Notes: The following three note fields (530, 533 and 535) relate to material available in a facsimile reproduction, such as microfilm. (See field 581 for published texts of documents).

530 Additional physical form available note

Indicates that the material described exists in the repository in the original form and as a reproduction. Subfields include the type of reproduction available (e.g., microfilm or microfiche), the source, the availability condition (e.g., for sale), and the order number.

533 Reproduction note

Notes that the archives has only a reproduction of the material being described (e.g., accessioned microfilm).

535 Location of originals/duplicates

Indicates the location of the originals if the repository creating the MARC-AMC record has only a reproduction. It can also be used to indicate the location of additional copies (such as in regional branches) of microfilmed records.

540 Terms governing use and reproduction

Indicates that a restriction exists on the use of records. For example, a film may have no access restriction and therefore be viewable. A use restriction, however, may prohibit reproduction without permission for access restrictions, see field 506 above).

541 Immediate source of acquisition

Notes the source for the material being described.

544 Location of associated materials

Contains a brief note on material that has the same provenance as the material being described but that is located in another repository.

.../
545 Biographical or historical note
Contains information about the person or organization that created the materials being described.

546 Language note
Notes the language or languages in which the records being described are written.

555 Cumulative index/finding aids note
Identifies indexes and finding aids available for the records being described.

561 Provenance
Consists of a narrative discussion of the provenance or custodial history of the records being described.

565 Case file characteristics note
Consists of statistical information about case files with subfields for number of cases/variables, names of variables, unit of analysis, universe of data, and filing scheme or code.

581 Publications note
Notes that the material being described has been published in a book or journal. The MARC-AMC record containing a description of the Department of State decimal file could have the entry: "Significant documents in this series have been published in the Foreign Relations of the United States."

583 Actions
Notes actions that have been taken on the records or parts of the records being described.

584 Accumulation and frequency of use
Notes the expected accumulation of the records being described and how frequently they are expected to be used by researchers.

.../
59X Local notes

These 10 fields (590-599) may be defined locally.

Subject Added Entries (6XX) and Added Entries (7XX)

In library systems, subject added entries and added entries were used originally to produce additional cards, one card to be filed for each additional entry. Subject added entries, the 6XX fields, contain descriptive information about the records. Added entries, the 7XX fields, generally contain additional information about the creators of the records (such as the name of an additional agency that created materials).

In archival indexing, the distinction between "by" and "about" usually is not defined as clearly as it is in cataloging books. Records or manuscripts are generally both by and about their creators. A book is generally by an author but not about the same person. Only an autobiography is both by and about an author. Because some archivists believe that the distinction is not worth maintaining, many archives rely more heavily on the 6XX fields than on the 7XX fields when a parallel exists.

Most 6XX and 7XX fields are entered under authority control to improve retrieval (i.e., the terms used in the fields come from a standardized list). Most institutions generally use the Library of Congress subject headings and name authority file.

600 Subject added entry - personal name

Identifies a person who is the subject of the records.

610 Subject added entry - corporate name

Identifies a corporation, including a government agency, that is the subject of the records.

650 Subject added entry - topical heading

Indicates subjects that are identified in the records.

651 Subject added entry - geographic name

Lists geographic terms describing the content of records.
655 Genre/form heading

Identifies the types of records being described, such as logbooks, diaries, or census schedules.

656 Index term - occupation

Identifies the occupation(s) of the records creator or the occupation(s) identified in the record.

657 Index term - function

Identifies the governmental functions documented in the records. Currently, there is no standard list of terms from which to choose. Seven state archives, as part of a National Historical Publications and Records Commission grant, are preparing a preliminary list for state governments.

700 Added entry - personal name

710 Added entry - corporate name

711 Added entry - conference or meeting

773 Host item entry

Identifies the "superior" record in a data base. For example, it can be used to identify the subgroup of which a series description is part. Subfields can identify the main entry, title, and identification number of the superior record.

Other Fields (8XX)

The remainder of the fields contain a variety of information not easily categorized.

851 Location

Contains the mailing address of the institution holding the materials described.

.../
Local fields (9XX)

Fields 900 through 999 may be defined and subdivided as required by the local institution. It is doubtful that such information could be accepted and used by an external system.
<table>
<thead>
<tr>
<th>Tag</th>
<th>Field Title</th>
<th>Tag</th>
<th>Field Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>001</td>
<td>Control number</td>
<td>520</td>
<td>Summary, abstract, annotation, scope etc., note</td>
</tr>
<tr>
<td>002</td>
<td>Subrecord map of directory</td>
<td>521</td>
<td>Users/intended audience note</td>
</tr>
<tr>
<td>005</td>
<td>Date and time of latest transaction</td>
<td>524</td>
<td>Preferred citation of described materials</td>
</tr>
<tr>
<td>007/00</td>
<td>Category of material</td>
<td>530</td>
<td>Additional physical form available note</td>
</tr>
<tr>
<td>007/01</td>
<td>Specific material designation</td>
<td>533</td>
<td>Reproduction note</td>
</tr>
<tr>
<td>007/02</td>
<td>Original versus reproduction aspect</td>
<td>535</td>
<td>Location of originals/duplicates</td>
</tr>
<tr>
<td>007/03</td>
<td>Polarity (microforms)</td>
<td>540</td>
<td>Terms governing use and reproduction</td>
</tr>
<tr>
<td>007/04</td>
<td>Dimensions (microforms)</td>
<td>541</td>
<td>Immediate source of acquisition</td>
</tr>
<tr>
<td>007/05- 08</td>
<td>Reproduction ratio</td>
<td>542</td>
<td>Location of associated materials</td>
</tr>
<tr>
<td>007/09</td>
<td>Color (microforms)</td>
<td>543</td>
<td>Biographical or historical note</td>
</tr>
<tr>
<td>007/10</td>
<td>Emulsion on film (microforms)</td>
<td>546</td>
<td>Language note</td>
</tr>
<tr>
<td>007/11</td>
<td>Generation</td>
<td>555</td>
<td>Cumulative index/finding aids note</td>
</tr>
<tr>
<td>007/12</td>
<td>Base of film (microforms)</td>
<td></td>
<td>Provenance</td>
</tr>
<tr>
<td>008/00- 05</td>
<td>Date entered on file</td>
<td>561</td>
<td>Copy and version identification</td>
</tr>
<tr>
<td>008/06</td>
<td>Type of date code</td>
<td>562</td>
<td>Case file characteristics note</td>
</tr>
<tr>
<td>008/07- 10</td>
<td>Date 1</td>
<td>565</td>
<td>Linking entry complexity note</td>
</tr>
<tr>
<td>008/11- 14</td>
<td>Date 2</td>
<td>580</td>
<td>Publications note</td>
</tr>
<tr>
<td>008/15- 17</td>
<td>Place of publication, production, or execution code</td>
<td>584</td>
<td>Actions</td>
</tr>
<tr>
<td>008/18- 19</td>
<td>Local notes</td>
<td>59X</td>
<td>Accumulation and Frequency of use</td>
</tr>
<tr>
<td>008/23</td>
<td>Form of reproduction code</td>
<td>600</td>
<td>Subject added entry - personal name</td>
</tr>
<tr>
<td>008/24- 22</td>
<td>Undefined</td>
<td>610</td>
<td>Subject added entry - corporate name</td>
</tr>
<tr>
<td>008/35- 34</td>
<td>Undefined</td>
<td>611</td>
<td>Subject added entry - conference or meeting</td>
</tr>
<tr>
<td>008/37</td>
<td>Language code</td>
<td>630</td>
<td>Subject added entry - uniform title heading</td>
</tr>
<tr>
<td>008/38</td>
<td>Modified record code</td>
<td>650</td>
<td>Subject added entry - topical heading</td>
</tr>
<tr>
<td>008/39</td>
<td>Cataloging source code</td>
<td>651</td>
<td>Subject added entry - geographic name</td>
</tr>
<tr>
<td>010</td>
<td>Library of Congress control number</td>
<td>655</td>
<td>Genre/form heading</td>
</tr>
<tr>
<td>035</td>
<td>Local system control number</td>
<td>656</td>
<td>Index term - occupation</td>
</tr>
<tr>
<td>039</td>
<td>Level of bibliographic control and coding detail</td>
<td>657</td>
<td>Index term - function</td>
</tr>
<tr>
<td>040</td>
<td>Cataloging source</td>
<td>69X</td>
<td>Local subject added entries</td>
</tr>
<tr>
<td>041</td>
<td>Language code</td>
<td>700</td>
<td>Added entry - personal name</td>
</tr>
<tr>
<td>043</td>
<td>Geographic area code</td>
<td>710</td>
<td>Added entry - corporate name</td>
</tr>
<tr>
<td>045</td>
<td>Chronological code or date/time</td>
<td>711</td>
<td>Added entry - conference or meeting</td>
</tr>
<tr>
<td>052</td>
<td>Geographic classification code</td>
<td></td>
<td>Added entry - uniform title heading</td>
</tr>
<tr>
<td>066</td>
<td>Character sets present</td>
<td>730</td>
<td></td>
</tr>
<tr>
<td>072</td>
<td>Subject category code</td>
<td></td>
<td></td>
</tr>
<tr>
<td>09X</td>
<td>Local call numbers</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## AMC FORMAT VARIABLE DATA FIELDS (cont'd)

<table>
<thead>
<tr>
<th>Tag</th>
<th>Field Title</th>
<th>Tag</th>
<th>Field Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>Main entry - personal name</td>
<td>740</td>
<td>Added entry - title traced</td>
</tr>
<tr>
<td>110</td>
<td>Main entry - corporate name</td>
<td>752</td>
<td>Added entry - place of publication or production</td>
</tr>
<tr>
<td>111</td>
<td>Main entry - conference or</td>
<td>773</td>
<td>Host item entry</td>
</tr>
<tr>
<td>130</td>
<td>Main entry - uniform title heading</td>
<td>851</td>
<td>Location</td>
</tr>
<tr>
<td></td>
<td></td>
<td>870</td>
<td>Variant personal name</td>
</tr>
<tr>
<td></td>
<td></td>
<td>871</td>
<td>Variant corporate name</td>
</tr>
<tr>
<td>240</td>
<td>Uniform title</td>
<td>872</td>
<td>Variant conference or meeting name</td>
</tr>
<tr>
<td>242</td>
<td>Translation of title by cataloging agency</td>
<td>873</td>
<td>Variant uniform title heading</td>
</tr>
<tr>
<td>243</td>
<td>Uniform title, collective</td>
<td>880</td>
<td>Alternate graphic representation</td>
</tr>
<tr>
<td>245</td>
<td>Title statement</td>
<td>886</td>
<td>Foreign MARC information field</td>
</tr>
<tr>
<td>260</td>
<td>Publication, distribution, etc.,</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(imprint)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>300</td>
<td>Physical description</td>
<td></td>
<td></td>
</tr>
<tr>
<td>340</td>
<td>Medium</td>
<td></td>
<td></td>
</tr>
<tr>
<td>351</td>
<td>Orgination and arrangement</td>
<td></td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>General note</td>
<td></td>
<td></td>
</tr>
<tr>
<td>502</td>
<td>Dissertation not</td>
<td></td>
<td></td>
</tr>
<tr>
<td>505</td>
<td>Contents note (formatted)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>506</td>
<td>Restrictions on access</td>
<td></td>
<td></td>
</tr>
<tr>
<td>510</td>
<td>Citation note (brief form) references</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
INTERNATIONAL COUNCIL ON ARCHIVES

Section of Archivists of International Organizations (SIO)

THIRTEENTH SESSION

(IMF, Washington, D.C., 7-9 October 1987)

WORKING PAPER ON ITEM 4 OF THE PROVISIONAL AGENDA

The Computerized Archive Management and Retrieval System in the CERN Archive

Summary

In 1979 the decision has been taken to create the CERN Historical Archive after 25 years of existence of the Organization and to use INFOL, a readily available computer system, for the cataloguing and data retrieval. INFOL is working at CERN on the central IBM computer. The application of INFOL-2 for the CERN Archive inventory and the search procedures are described.

General remarks:

CERN is an international governmental organization created in 1954 and has now about 5000 members of staff. One of the particularities of its administrative structure is that it has not got a central registry, nor a central records management system.

The database of the CERN Historical Archive, established in 1980, comprises nowadays about 11 500 records.
The computerized archive management and retrieval system in the CERN archive

The decision to create the CERN Historical Archive as part of the Scientific Information Service has been taken in 1979 and it was decided to compile the inventory on a computer data base. At that time the only readily available system on a central computer at CERN with relatively simple access was INFOL, invented at the University of Geneva, in collaboration with the Data Handling Division of CERN.

INFOL seemed to fulfill our requirements in the archive. INFOL stands for INFORMATION Oriented Language. It was originally created for the CDC 3600/3800 series computers. INFOL-2, the present version is mainly written in ANSI FORTRAN IV and is available from the Centre Interfacultaire de Calcul Electronique (CICE) at the University of Geneva. It is currently implemented at some 30 sites on various hardware. Versions are available for CDC, IBM, UNIVAC, CII-HB and PDP-10.

1. Definitions

In the terminology of INFOL we define:

1.1 FILE as the collection of information to be handled by INFOL

1.2 ELEMENT as a unit of information often known as "record". A FILE consists of a sequentially ordered set of ELEMENTS

1.3 ITEM is part of an ELEMENT, usually known as "field".

2. Phases

INFOL is divided into 4 main phases:

2.1 ESTABLISHMENT to create the file

2.2 INTERROGATION for information retrieval

2.3 UPDATE to add, remove or modify elements of the file

2.4 REVISION allows partial modification of the specifications of an INFOL file.
3. **Description of the Archive Inventory**

After having given the above definitions it might be easier to explain further details of the system using INFOL in our archive inventory.

3.1 **Structure of the Database** (see table 1)

3.1.1 We have structured our records (called "ELEMENTS" in INFOL) into 18 fields (called "ITEMS"). They are numbered from 1 to 18 as shown in the first column of the table. Each ITEM has a name (called "DESCRIPTION") given in the second column. The third column gives some further explanation where needed, but this does not belong to INFOL. The next column shows the CATEGORY and TYPE of every ITEM.

3.1.2 CATEGORY may be unary (no "sub-fields") or multiple (with two or more "sub-fields").

Remark: The serial number of an ELEMENT must always have ITEM-number 1 and is, by definition, a unary ITEM.

3.1.3 The following TYPES can be distinguished for an ITEM:

---

**ALPHANUMERIC ITEMS** may contain any character, except the asterisk which has a special function as delimiter. The length is limited to 700 characters; spaces are significant.

**NUMERIC ITEMS** contain only figures which may be integers or reals, positive or negative. Dates are internally converted into the form yy mm dd.

**CODED ITEMS** may only contain predefined codes, either numeric codes with consecutive figures from 1 to x or mnemonic codes, e.g. f for free, r for restricted and c for confidential (see ITEM *4* in table 1).

3.1.4 In each main phase of INFOL, the system will check CATEGORY and TYPE of every ITEM and reject input when there is an error. Additional criteria for checking input are to be defined in VALIDATION (see last column of table 1). An ITEM may be necessary which means that if it is not present the input will be rejected.

3.1.5 The maximum number of characters in an ITEM or SUB-ITEM is defined in a parameter called "CHARACTERS" whereas for a multiple ITEM the number of possible SUB-ITEMS must be defined in a parameter called "MAXIMUM". A numeric ITEM may be defined as "INTEGER".
3.2 Input of information for the Archive Inventory

3.2.1 The input of information must be given in a structured form with an asterisk around the ITEM-numbers and an additional asterisk at the end of every ELEMENT. SUB-ITEMS in multiple ITEMS are separated by an asterisk. When entered in a string, the information must be placed into 80 columns.

3.2.2 As this input-method is quite cumbersome it is more convenient to use a structured auxiliary file called WYLBUR. Table 2 shows the structure of our auxiliary input file and table 3 part of an input file.

3.2.3 The UPDATE phase is quite similar to the ESTABLISHMENT phase. If it becomes necessary to change the structure of an INFOL file this can be done, with certain limitations, in a REVISION phase which can also be used to add new codes.

4. The Interrogation of the Archive Inventory

To make a search in the Archive data bases specific RETRIEVAL criteria must be met. The search is sequential. There are no inverted files. We do not work with a thesaurus of keywords but the subject approach is possible by the search of free text in the ITEMS*Title* and *Title augmentation*.

Several RETRIEVAL criteria may be formulated in one job by specifying RETR CRIT n where n is a natural figure assigned in sequence.

4.1 RETRIEVAL criteria

There are existence criteria and relational criteria:

4.1.1 Existence criteria:
EXISTS means that the corresponding ITEM must not be empty (e.g. *1* EXISTS calls all documents in the file).
DOES NOT EXIST searches for all ELEMENTS in which the corresponding ITEM is empty.

4.1.2 Relational criteria:
Two logical operators are available in INFOL to link two or several search criteria
AND expresses the logical product
OR expresses the logical sum

\[ \text{AND} \quad \text{OR} \]
The logical operator NOT does not exist on its own, but can be replaced by using complementary operators:

- equal EQ
- not equal NE
- less equal LE
- greater equal GE
- less than LT
- greater than GT

Note: In INFOL the logical operators AND and OR have the same priority and the logical combination will proceed from left to right. Parenthesis must be used if a different combination is meant.

Example: Search for "programming system" and "programming language" in ITEM 7 must be formulated:
*7* EQ *Programming* AND (EQ *System* OR EQ *Language*).

4.2 Extraction

The output of the result of a search is obtained according to specifications given after the control word EXTRACTION. The automatic report generation which is used in most cases, is easy to handle. All that is needed is the command REPORT after each ITEM number which should be printed or the command REPORT ALL if all ITEMS have to be printed.

5. Conclusion

INFOL is a retrieval language of an older generation which however well fulfills its function in the CERN Archive.

Searches are performed by submission of batch jobs from a terminal. Response time is of the order of 1-2 minutes for a data base of more than 11 thousand items.

I should like to conclude with a few examples of retrieval from our data base. (tables 3a, 3b)
<table>
<thead>
<tr>
<th>ITEM-No</th>
<th>ITEM-DESCRIPTION</th>
<th>EXPLANATION</th>
<th>CATEGORY-TYPE</th>
<th>VALIDATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>serial-no.</td>
<td></td>
<td>UNAR NUME</td>
<td>INTE</td>
</tr>
<tr>
<td>2</td>
<td>subject-category</td>
<td></td>
<td>MULT ALPH</td>
<td>CHAR 20 MAXI 3</td>
</tr>
<tr>
<td>3</td>
<td>box-no.</td>
<td>first and last box-no.</td>
<td>MULT ALPH</td>
<td>NECE CHAR 4 MAXI 2</td>
</tr>
<tr>
<td>4</td>
<td>collation</td>
<td>number of pages or thickness in cm</td>
<td>UNAR ALPH</td>
<td>CHAR 25</td>
</tr>
<tr>
<td>5</td>
<td>access-code</td>
<td>free, restricted or confidential</td>
<td>UNAR CODE</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>user-code</td>
<td>history project or archive</td>
<td>UNAR CODE</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>file-no</td>
<td>original document or file number</td>
<td>MULT ALPH</td>
<td>CHAR 30 MAXI 3</td>
</tr>
<tr>
<td>8</td>
<td>related file-no.</td>
<td>secondary numbers referred to on title page</td>
<td>MULT ALPH</td>
<td>CHAR 40 MAXI 3</td>
</tr>
<tr>
<td>9</td>
<td>compilers</td>
<td></td>
<td>MULT ALPH</td>
<td>CHAR 40 MAXI 6</td>
</tr>
<tr>
<td>10</td>
<td>pers.name-title</td>
<td></td>
<td>MULT ALPH</td>
<td>CHAR 40 MAXI 7</td>
</tr>
<tr>
<td>11</td>
<td>subject-title</td>
<td></td>
<td>UNAR ALPH</td>
<td>NECE CHAR 240</td>
</tr>
<tr>
<td>12</td>
<td>title-augmentation</td>
<td>keywords or phrases to improve retrieval</td>
<td>UNAR ALPH</td>
<td>CHAR 480</td>
</tr>
<tr>
<td>13</td>
<td>observations</td>
<td></td>
<td>UNAR ALPH</td>
<td>CHAR 480</td>
</tr>
<tr>
<td>14</td>
<td>dates of file</td>
<td>opening and closing dates of the file</td>
<td>MULT DATE</td>
<td>NECE MAXI 2</td>
</tr>
<tr>
<td>15</td>
<td>reception date</td>
<td>date of reception in the archive</td>
<td>UNAR DATE</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>technical support</td>
<td>only given if non-standard (e.g. tape)</td>
<td>UNAR ALPH</td>
<td>CHAR 20</td>
</tr>
<tr>
<td>17</td>
<td>person responsible</td>
<td>in general: person who deposited-the file</td>
<td>UNAR ALPH</td>
<td>CHAR 80</td>
</tr>
<tr>
<td>18</td>
<td>housekeeping</td>
<td>notes related to archive administration</td>
<td>UNAR ALPH</td>
<td>CHAR 240</td>
</tr>
</tbody>
</table>

CODES:  
4 mnemonic f free r restricted c confidential  
5 mnemonic a archive h history-study
<table>
<thead>
<tr>
<th>Column</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>serial-no.</td>
</tr>
<tr>
<td>2</td>
<td>subject-category (maxi 3) (maxi 20 char. each)</td>
</tr>
<tr>
<td></td>
<td>category 1 =&gt;</td>
</tr>
<tr>
<td></td>
<td>category 2 =&gt;</td>
</tr>
<tr>
<td></td>
<td>category 3 =&gt;</td>
</tr>
<tr>
<td>3</td>
<td>box-no. (maxi 4 char. each)</td>
</tr>
<tr>
<td></td>
<td>box (first) =&gt;</td>
</tr>
<tr>
<td></td>
<td>box (last) =&gt;</td>
</tr>
<tr>
<td>4</td>
<td>collation (e.g. number of pages) (maxi 25 char.) =&gt;</td>
</tr>
<tr>
<td>5</td>
<td>access-code (f, r or c) (maxi 1 char.) =&gt;</td>
</tr>
<tr>
<td>6</td>
<td>user-code (a or h) (maxi 1 char.) =&gt;</td>
</tr>
<tr>
<td>7</td>
<td>file-no. (maxi 3) (maxi 30 char. each)</td>
</tr>
<tr>
<td></td>
<td>file-no. 1 =&gt;</td>
</tr>
<tr>
<td></td>
<td>file-no. 2 =&gt;</td>
</tr>
<tr>
<td></td>
<td>file no. 3 =&gt;</td>
</tr>
<tr>
<td>8</td>
<td>related-file no. (maxi 3) (maxi 40 char. each)</td>
</tr>
<tr>
<td></td>
<td>related-file no. 1 =&gt;</td>
</tr>
<tr>
<td></td>
<td>related-file no. 2 =&gt;</td>
</tr>
<tr>
<td></td>
<td>related-file no. 3 =&gt;</td>
</tr>
<tr>
<td>9</td>
<td>compilers (maxi 6) (maxi 40 char. each)</td>
</tr>
<tr>
<td></td>
<td>compiler 1 =&gt;</td>
</tr>
<tr>
<td></td>
<td>compiler 2 =&gt;</td>
</tr>
<tr>
<td></td>
<td>compiler 3 =&gt;</td>
</tr>
<tr>
<td></td>
<td>compiler 4 =&gt;</td>
</tr>
<tr>
<td></td>
<td>compiler 5 =&gt;</td>
</tr>
<tr>
<td></td>
<td>compiler 6 =&gt;</td>
</tr>
<tr>
<td>10</td>
<td>personal name-title (maxi 7) (maxi 40 char. each)</td>
</tr>
<tr>
<td></td>
<td>personal name 1 =&gt;</td>
</tr>
<tr>
<td></td>
<td>personal name 2 =&gt;</td>
</tr>
<tr>
<td></td>
<td>personal name 3 =&gt;</td>
</tr>
<tr>
<td></td>
<td>personal name 4 =&gt;</td>
</tr>
</tbody>
</table>
personal name 5 ==> 
personal name 6 ==> 
personal name 7 ==> 

*11* subject-title (maxi 240 char.) ==> 

*12* title-augmentation (maxi 480 char.) ==> 

*13* observations (maxi 480 char.) ==> 

*14* dates of files (format yymmmdd) 
  opening date ==> 
  closing date ==> 

*15* reception date (format yymmmdd) ==> 

*16* technical support (e.g. paper) (maxi 20 char.) ==> 

*17* person responsible (maxi 80 char.) ==> 

*18* housekeeping (maxi 240 char.) ==>
NUMBER OF ELEMENTS WHICH SATISFY THE SEARCH = 13 (0.41 % OF THE FILE)

DETAILS =

<table>
<thead>
<tr>
<th>I</th>
<th>SPECIFIC DETAILS OF THE SEARCH</th>
<th>I NUMBER OF ELEMENTS</th>
<th>I PERCENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td></td>
<td>1138</td>
<td>100.00</td>
</tr>
<tr>
<td>I</td>
<td>AND 11x EXIST</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>OR 12x EXIST</td>
<td>15</td>
<td>0.48</td>
</tr>
<tr>
<td>I</td>
<td>AND 14x GE 19600000</td>
<td>25</td>
<td>0.80</td>
</tr>
<tr>
<td>I</td>
<td>AND 14x LE 19870000</td>
<td>13</td>
<td>0.41</td>
</tr>
</tbody>
</table>

--ELAPSED TIME FOR RETRIEVAL TASK HR. = 0.594 SECS.

REPORT ALL

Table 3a
<table>
<thead>
<tr>
<th>serial-no.</th>
<th>20081</th>
</tr>
</thead>
<tbody>
<tr>
<td>subject-category</td>
<td>DIR/ADM Files</td>
</tr>
<tr>
<td>box-no.</td>
<td>F211</td>
</tr>
<tr>
<td>collation</td>
<td>Classeur Federal/6 cm</td>
</tr>
<tr>
<td>access-code</td>
<td>r</td>
</tr>
<tr>
<td>user-code</td>
<td>h</td>
</tr>
<tr>
<td>file-no</td>
<td>DIR/ADM/15 (1)</td>
</tr>
<tr>
<td>subject-title</td>
<td>Droits d'auteurs.</td>
</tr>
<tr>
<td>observations</td>
<td>() Convention de Berne, rev. 1948.</td>
</tr>
<tr>
<td>dates of file</td>
<td>MAR/1949 # APR/1975</td>
</tr>
<tr>
<td>reception date</td>
<td>JAN/1981</td>
</tr>
<tr>
<td>technical support</td>
<td>paper</td>
</tr>
<tr>
<td>person responsible</td>
<td>S. Tracy</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>serial-no.</th>
<th>20324</th>
</tr>
</thead>
<tbody>
<tr>
<td>subject-category</td>
<td>DIR/ADM Files</td>
</tr>
<tr>
<td>box-no.</td>
<td>F530</td>
</tr>
<tr>
<td>collation</td>
<td>Classeur Federal/8 cm</td>
</tr>
<tr>
<td>access-code</td>
<td>r</td>
</tr>
<tr>
<td>user-code</td>
<td>h</td>
</tr>
<tr>
<td>file-no</td>
<td>DIR/ADM/DD/01 # File 1</td>
</tr>
<tr>
<td>subject-title</td>
<td>Data Handling Division (DD): General.</td>
</tr>
<tr>
<td>title-augmentation</td>
<td>Jungfrevich equipment 1958/1959. CERN/STS/Memo on electronic instrumentation. Research activities in STS Division. Reports on computer facilities at CERN and future computing needs, as well as Finance Committee and Scientific Policy Committee documents on the subject, etc. Decisions of Finance Committee.</td>
</tr>
<tr>
<td>observations</td>
<td>File in bad condition (flood damage). DD Division as from 1960, 1955-1959 STS Division (Scientific and Technical Services Division).</td>
</tr>
<tr>
<td>dates of file</td>
<td>OCT/1955 # DEC/1966</td>
</tr>
<tr>
<td>reception date</td>
<td>JAN/1981</td>
</tr>
<tr>
<td>technical support</td>
<td>paper</td>
</tr>
<tr>
<td>person responsible</td>
<td>S. Tracy</td>
</tr>
</tbody>
</table>

and so on
--ELAPSED TIME FOR EXTRACTION SET NR.  1 ...  0.060 SECS.

----TOTAL TIME FOR RETRIEVAL SET NR.  1 ...  0.654 SECS.

--INFOL EXECUTION TERMINATED NORMALLY.
**INFORMATION RETRIEVAL REPORT NO. 1**

**INTE RETR CRIT 1**

*1* EXIST

AND (*11* EQ #bubble# AND ( EQ #detector# OR EQ #counter# ))

OR (*12* EQ #bubble# AND ( EQ #detector# OR EQ #counter# ))

*14* GE 1957

*14* LE 1987

**EXTR 1**

---------------------------------------------------------------------

**STATISTICAL ANALYSIS**

---------------------------------------------------------------------

**NUMBER OF ELEMENTS WHICH SATISFY THE SEARCH = 3 ( 0.04 0% OF THE FILE)**

**DETAILS =**

---------------------------------------------------------------------

<table>
<thead>
<tr>
<th>I</th>
<th>SPECIFIC DETAILS OF THE SEARCH</th>
<th>I</th>
<th>NUMBER OF ELEMENTS</th>
<th>I</th>
<th>PERCENT I</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>I WHICH SATISFY EACH I</td>
<td>I</td>
<td>CRITERION</td>
<td>I</td>
<td>I</td>
</tr>
<tr>
<td>I</td>
<td>I</td>
<td>I</td>
<td>I</td>
<td>I</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>I</th>
<th>I</th>
<th>I</th>
<th>I</th>
</tr>
</thead>
<tbody>
<tr>
<td>AND</td>
<td>11</td>
<td>EXIST</td>
<td>7618</td>
</tr>
<tr>
<td>I</td>
<td>AND</td>
<td>11</td>
<td>EQ BUBBLE</td>
</tr>
<tr>
<td>I</td>
<td>AND</td>
<td>11</td>
<td>EQ DETECTOR</td>
</tr>
<tr>
<td>I</td>
<td>OR</td>
<td>11</td>
<td>EQ COUNTER</td>
</tr>
<tr>
<td>I</td>
<td>OR</td>
<td>12</td>
<td>EQ BUBBLE</td>
</tr>
<tr>
<td>I</td>
<td>AND</td>
<td>12</td>
<td>EQ DETECTOR</td>
</tr>
<tr>
<td>I</td>
<td>OR</td>
<td>12</td>
<td>EQ COUNTER</td>
</tr>
<tr>
<td>I</td>
<td>AND</td>
<td>14</td>
<td>GE 19570000</td>
</tr>
<tr>
<td>I</td>
<td>AND</td>
<td>14</td>
<td>LE 19870000</td>
</tr>
</tbody>
</table>

**--ELAPSED TIME FOR RETRIEVAL TASK NR. 1... 2.502 SECS.**

**REPORT ALL**

---

Table 3b
| serial-no. | 108 |
| box-no.    | A234 |
| collation  | 5 p |
| file-no    | CERN/SPC/51 |
| subject-title | 10th Meeting of the Scientific Policy Committee, 4 September 1958. Proposals on PS Experimental Programme. |
| dates of file | 01/SEP/1958 |

| serial-no. | 130 |
| box-no.    | A235 |
| collation  | 1 p + Annexes (35 p) |
| file-no    | CERN/SPC/90 |
| subject-title | 12th Meeting of the Scientific Policy Committee, 17 April 1959. PS Experimental Programme. |
| title-augmentation | Construction programme. The long flight path. Cloud chamber. Hydrogen bubble chamber. IEP activities. Gas & other Cerenkov counters. Emulsions techniques, etc. |
| dates of file | 31/MAR/1959 |

| serial-no. | 188 |
| box-no.    | A235 |
| collation  | 9 p + Annex (2 p) |
| file-no    | CERN/SPC/127/Rev. |
| title-augmentation | Ecole Polytechnique bubble chamber. The counter & cloud chamber set-up. The CERN bubble chamber, etc. |
| observations | Annex: drawings. |
| dates of file | 18/APR/1961 |
---ELAPSED TIME FOR EXTRACTION SET NR. 1 ... 0.020 SECS.

----TOTAL TIME FOR RETRIEVAL SET NR. 1 ... 2.522 SECS.

--INFOL EXECUTION TERMINATED NORMALLY.
A brief review of the development of Registry systems suggests that "modernization" has usually meant a move towards increased complexity in document handling and arrangement to the detriment of specificity of indexing. Automation makes feasible a return to document-level control with identification of documents in an electronic filing system supplied by the system and the originator. Arrangement of file folders in a subject classification scheme should be abandoned.
1. Is the Registry system an old-fashioned, inefficient method of dealing with correspondence or is it the archetype of a modern, fully automated, electronic filing system? This is an intriguing question, of more than academic interest to one who has the mandate to "modernize" the registry system of a forty year old international organization. A brief review of the development of Registry systems suggests that "modernization" has usually meant a move towards increased complexity in document handling and arrangement, to facilitate the existence of users, and to the detriment of specificity of indexing. It is debatable whether such changes which fall under the description of modernization have really contributed to increased effectiveness of the system. The challenge is to identify the features of registry systems which are most productive in achieving the kind of control needed by modern organizations and focusing our planning on those features.

2. Theodore R. Schellenberg gives a brief history of the origins of registry systems and of their development in the national traditions of Germany, England and Australia. Early registries maintained records in two chronologically-arranged series of documents - incoming and outgoing - entered into a register and numbered consecutively with index references to subjects and persons keyed to the documents. Assembling the individual items or tracing through the index to consult all the documents relevant to a subject, event, person, issue to provide the "complete story" aimed at by modern separate files would obviously have been time-consuming, especially as the volume of records increased. Presumably that is one reason for the development of an enhancement in Registry systems - the separate or individual file unit.

3. In more advanced registry systems, file folders established on the theory of one subject one file bring together both incoming and outgoing documents in one series. If the register continues to be used for individual documents, the file number is added as a location. In some cases the Register becomes a register of file folders rather than individual documents. Index references to the name of writers and subjects of documents are keyed to the file folders, and (in some cases) to the "folio" number of the document within the file.

4. The classified Registry represents another level of complexity or simplicity - the decision to place the folders in order by subject rather than by a number assigned in the consecutive order of accumulation. The file folders are arranged in a classification scheme with main headings and sub-headings, and subject indexes are keyed to headings in the Classification scheme rather than to the file folder or the individual document.

5. Thus in the progressively more structured Registries, we often see a consecutive shift in the indexing target from individual documents, to the file folders, to the subject classification scheme itself.

6. In the Australian single number system, as described by Schellenberg, documents are grouped in file folders which are numbered consecutively in order of accumulation. Subject access is through the subject index - an authoritative list of indexable headings. Subject references for each file can be added to or modified as the contents of the file accumulate without changing the location of the files.

7. In summary, although there are as many Registry systems as organizations with registries, the general characteristics of a Registry system include the following:

- Registration of documents: on an inclusive or selective basis;

- File folders established on the basis of one subject one file and consisting of a collection of pieces concerning an affair, a specific person or organization, topic, event, concept, location;

- File arrangement: Either in a simple consecutive numerical sequence or in accordance with a classification scheme;

- Subject access: Either through arrangement of the physical files in a classification scheme or by a subject index to file folders;

- Aids to locate files/documents: Indexes to persons, subjects, meetings, etc., or lists or classification manuals.

8. The registry system of WHO is typical in that it is a hybrid system which has evolved in the forty years of utilization. There is registration of individual documents, but only on a highly selective basis - about 10% of documents handled by Registry or 2% of the total correspondence of the Organization. File folders are established on the basis of one subject one file and the file folders are classified and physically arranged according to a classification scheme which features primary subject headings sub-divided by standard secondary headings. Subject indexing is provided by the classification system itself. Indexing by originator organizations is done on a highly selective basis.

9. Long-term members of Registry staff asserts "simplification projects" have resulted in a steady erosion of the effectiveness of Registry. Such projects have only reduced the amount of registration and indexing of individual documents. In fact, the time gained by recording and indexing only a percentage of documents, is lost because of increased time necessary to retrieve and the vicious cycle continues. Reduced efficiency in retrieval results in an undermining of organizational confidence in the service, and an increase in duplicate files, which increases the workload in the Archives.
10. Clearly automation has great potential to improve the management of correspondence and Registry procedures. Such means give us the opportunity to make current procedures more efficient, to improve the quality of services of Registry, and even, perhaps, to do what we have always said we should have done or would like to do—such as increase item-level indexing. The computer presents us with a great number of choices and with alternate ways of achieving ends. It should also result in a thorough evaluation of our procedures; how we do things and why we do them.

11. Subject classification, for example. What is the purpose of arranging the file folders according to a subject classification scheme? Originally, I propose, it was to eliminate the need for subject indexing which was a labour intensive and specialized work, in other words: expensive. The second reason was that, it was thought, since a subject can only be handled by one service and at a certain level, the subject classification schemes would be able to eliminate duplicate files and get closer to the concept of provenance. If one understood the file scheme, one would know where to look for a file on a given subject and could browse in the same area to find related files. If the file codes were mnemonic, classification made the file codes easier to remember than consecutive meaningless numbers. The classification number became a shorthand for the file content.

12. If a computer system is designed to provide direct subject indexing to the file folders, what is the purpose of arranging the files themselves in a subject order? Do users of the files browse the file storage area? It is easier for staff to assign file designations if the numbers mean something? My conclusion is that subject indexing makes subject arrangement, and probably subject codification, unnecessary.

13. Or the question of individual file units. File units were developed as organizational units multiplied, and the volume of documentation increased, in order to assemble the documentation in sequential order on particular issues and to provide a convenient unit for storage, transfer, and control. Again, filing together all the correspondence from a particular person/organization or concerning a specific case reduced the necessity for recording and indexing each item that fell within that category. The disadvantage is that while there may be "one subject-one file," a document may be analyzed as belonging in more than one file and a file as pertaining to more than one subject. Photocopying and cross-referencing might ease the situation but the file unit concept, if conceived as a physical location, is, unfortunately, both confining and still necessary. If, however, the file unit is seen as an intellectual grouping pertaining to more than one location, then the concept becomes very useful. Electronic filing systems, particularly optical disk, offer the possibility to assemble dossiers, on demand, according to criteria established by the user.
14. Automation can expand the usefulness of the file unit concept, but sophisticated indexing, manual or automated, becomes then a prerequisite. Until we have an electronic filing system, the file unit for physical storage and handling is necessary, but if we envisage a long-term move towards automation, then indexing and control at the file folder level only, will be inadequate and ineffective. Only if we control the documents themselves, will we be able to combine and recombine in dossiers according to users specifications.

15. If we conclude that item-level control is the most desirable method of handling documents, using the file unit as a flexible intellectual control concept rather than a physical file location, we must return to the question of why this level of control has had to be abandoned in most registries (except in the counter-intelligence registries described in the spy novels of John Le Carré). It was generally considered as too much work, too labour-consuming. Even with automation the problem remains. Data entry on an item level, unless very carefully planned, can make such a plan unfeasible.

16. However, systems can be designed to minimize the amount of data entry time by Registry staff and maximize the amount of identifying information that is supplied by the system (date, office of origin, author) or by the originator (destination, file category, subject). Charles Dollar has explored this approach in his RAMP study Electronic Records Management and Archives in International Organizations.

17. Conclusions:

As a starting point for discussion, I propose the following implications of automation for Registry systems:

1. Records managers must be involved in both short and long-range planning procedures. If we foresee an electronic filing system as our long-range goal, then we should plan our systems based on document level control.

2. In the short-term we will need physical units such as files to move documents around in, even if in the long run the file folder level is replaced by "search strategies" leading to the assembling of dossiers on demand.

3. The practice of arranging physical files according to a classification system should be abandoned and replaced by simple consecutive numbering of folders and an automated subject index.

4. Classification coding may be useful as a shorthand for the contents of a file or dossier, but the time and energy spent in maintaining classification systems would be better spent in improving subject indexing.
<table>
<thead>
<tr>
<th>Document ID</th>
<th>Date</th>
<th>Title</th>
<th>Author(s)</th>
<th>Version</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIO/87/WP-1</td>
<td>7.10.87</td>
<td>&quot;Implications of Automation for Registry Systems&quot;</td>
<td>Liisa FAGERLUND</td>
<td>E</td>
</tr>
<tr>
<td>SIO/87/WP-2</td>
<td>7.10.87</td>
<td>&quot;Report on Question of SIO Membership&quot;</td>
<td>Estanislao J. CIEŃEWČZ</td>
<td>E</td>
</tr>
<tr>
<td>SIO/87/WP-3</td>
<td>7.10.87</td>
<td>&quot;MARC-AMC Format&quot;</td>
<td>Alf M.E. ERLANDSSON</td>
<td>E</td>
</tr>
<tr>
<td>SIO/87/WP-4</td>
<td>7.10.87</td>
<td>&quot;SIO Five-Year Plan for Automation Activities&quot;</td>
<td>Gertrude D. LONG</td>
<td>E</td>
</tr>
<tr>
<td>SIO/87/WP-5</td>
<td>7.10.87</td>
<td>&quot;Cooperation Between SIO and the Advisory Committee for the Coordination of Information Systems (ACCIS) of the United Nations&quot;</td>
<td>Gertrude D. LONG</td>
<td>E</td>
</tr>
<tr>
<td>SIO/87/WP-6</td>
<td>7.10.87</td>
<td>&quot;Report on the Activities Following ACC Decision 1984/15&quot;</td>
<td>Gertrude D. LONG</td>
<td>E</td>
</tr>
<tr>
<td>SIO/87/WP-7</td>
<td>7.10.87</td>
<td>&quot;ICA Medium-Term Plan, 1988-1992&quot;</td>
<td></td>
<td>E</td>
</tr>
<tr>
<td>SIO/87/WP-8</td>
<td>7.10.87</td>
<td>&quot;The Computerized Archive Management and Retrieval System in CERN's Archives&quot;</td>
<td>Roswitha RAHMÖY</td>
<td>E</td>
</tr>
</tbody>
</table>
Partial List of MicroMARC:amc Purchasers

- Chicago Historical Society, Prints and Photographs Division
- University of Cincinnati, Archives and Special Collections
- Hallmark Cards, Historical Collection
- Historical Foundation of the Presbyterian and Reformed Churches, Archives
- Indiana Historical Society, Manuscripts Department
- Landsarkivet, Lund, Sweden
- Loma Linda University, University Archives
- University of Minnesota, Immigration History Research Center
- New Mexico State University, University Archives
- Northern Michigan University, University Archives
- City of Seattle, City Archives
- Texas A. & M. University, University Archives
- Western Michigan University, University Archives
- Oregon State University, Special Collections
- Ursuline Sisters, (Louisville, KY), Archives
- Archdiocese of Chicago, Archives and Records Center
- University of Michigan, Bentley Historical Library

For OCLC Subscribers

OCLC states: "We are able to read the MicroMARC:amc records, and... process them for loading into OCLC's Online Union Catalog... if the records pass a standard evaluation."

For More Information

Write to:

MicroMARC:amc
c/o University Archives and Historical Collections
EG-13 Library Building
Michigan State University
E. Lansing, MI 48824-1048
(517) 355-2330

Bitnet Address:
20669MMA@MSU

The MicroMARC:amc microcomputer system is a comprehensive system based on the USMARC AMC Format. With this system, you can enter records, update them with ease, and then produce full reports or do searches on your holdings. The system can also import or export a standard USMARC AMC formatted file. The system allows for maximum flexibility on a local basis while still adhering to the standards of the national format.
In response to the need for an improved records management system, the MicroMarc Corporation developed a new computer system named MicroMarc: America's Number One. This system is designed to assist in records management by providing a comprehensive set of features that allow users to efficiently manage their records. The MicroMarc: America's Number One system includes a wide range of features, such as the ability to create and edit records, search for records, and print reports. It is available in various configurations to suit the needs of different users, from small businesses to large organizations. The system is user-friendly and comes with detailed documentation to help users get up and running quickly. Whether you need to manage a small database or a large collection of records, MicroMarc: America's Number One is the perfect solution.
INTERNATIONAL COUNCIL
ON ARCHIVES

Section of Archivists of
International Organizations (SIO)

Thirteenth Session
Washington, D.C., IMF, 7, 8, and 9 October 1987

RECOMMENDATIONS

SIO/87/REC-1. Electronic Records Management and Archives
in International Organizations

Agenda Items 1, 3 and 5.1

A.

RECALLING the recommendation of Charles Dollar in his RAMP study
Electronic Records Management and Archives in International Organizations
that the Section of Archivists of International Organizations (SIO) of
ICA establish formal liaison with the United Nations Advisory Committee
for the Co-ordination of Information Systems (ACCIS) in order to ensure
the participation of archivists and records managers in the deliberations
of that body as they touch upon the vital interests and concerns of those
professions,

FURTHER RECALLING the discussions between the Executive Board of
SIO and the Executive Secretariat of ACCIS during the past year, which
culminated in the proposal for a Technical Panel on Records Management
which was submitted to the Advisory Committee for the Co-ordination
of Information Systems at its Fourth Session held in Geneva on
September 23-25, 1987,

NOTING with satisfaction the action taken at the Fourth Session
of ACCIS to establish such a Technical Panel, and

EXPRESSING its gratitude to the members of the Advisory Committee
for the Co-ordination of Information Systems for their interest and
support in this matter,

The Section of Archivists of International Organizations of the
International Council on Archives ENDORSES the ACCIS proposal that
there be a proper interfacing between the related activities of the
Technical Panel on Computer-Supported Telecommunication Services
(TP/COM) and the Technical Panel on Records Management, to avoid
duplication and conflicts of competence, and to this end,
URGES the Executive Secretariat of ACCIS to include in the membership of the Technical Panel on Records Management staff of member institutions who have experience in administering archives and records management programs. Such expertise would contribute to a thorough consideration of relevant issues and facilitate implementation of Panel recommendations. Such expertise would also enable the Panel to benefit from developments in the broader context of the archives and records management professions.

B.

Having discussed at length the issue of computer systems in archives and records management, as well as the more specific question of the implications of automation for registry systems, the Section of Archivists of International Organizations concludes that the development and implementation of international computer processing standards and protocols for the proper maintenance and preservation of electronic records is a matter of considerable priority. To this end, the Executive Board is directed to establish a close working relationship with the newly created ACCIS Technical Panel on Records Management, which has as one of its goals the development of standards for the maintenance and preservation of electronic mail, which standards could be accepted by the United Nations as system-wide standards and could become a part of X.400 as well. In its dealings with the Technical Panel, the Executive Board should at all times be aware of the usefulness of the Panel's work for non-UN system members of the Section, and should undertake to keep these members informed concerning developments.

SIO/87/REC-2. Optical Disk Technology

Agenda Item 2

The Section of Archivists of International Organizations of the International Council on Archives, noting that the Third Medium-Term Plan of ICA includes a provision for research on optical disks, directs its Executive Board to consult with the Automation Committee in this regard, with a view to co-sponsoring a study of available optical disk systems and evaluating them for use in archives and records management.

Agenda Item 4

The Executive Board is directed to submit to the International Council on Archives the following projects for inclusion in the ICA Third Medium-Term Plan (1988-1992):

1. Directory of Archives of International Organizations.

2. Survey of optical disc systems available for archives and records management applications, including evaluation of their suitability for use in an international environment. Support to be requested from UNESCO.

3. Establishment and maintenance of a clearinghouse for the exchange of information on software packages available for use in archives and records management applications.

4. Development of international computer processing standards for the maintenance of electronic records and their longer-term preservation and interpretation, (related to Project 4.2.1.2/CDP). Support expected from ACCIS.

The Executive Board is also requested to withdraw from the ICA Third Medium-Term Plan Item 0.2.2.9 (ICA/SIO Newsletter).

SIO/87/REC-4. Membership and Dues

Agenda Item 5.0

The Section of Archivists of International Organizations, of the International Council on Archives:

RECOGNIZING that it would be detrimental to its interests to exclude from participation in SIO those who are not members of ICA,

CONSIDERING that different levels of participation allow each participant to benefit from the Section's existence and enhance the individual and collective weight of its membership, while still permitting the fulfilling of individual needs,

RECONFIRMS that:

(a) A Member of the Section of Archivists of International Organizations is any archivist or records officer working for an international organization who is a Category "D" member of the International Council on Archives, or whose institution is a Category "C" member of that body.
Agenda Item 5.1

The Section of Archivists of International Organizations endorses the attached description of mission and goals relevant to the Five-Year Plan for Automation Activities, and directs the Planning Committee to prepare a draft plan to achieve the enumerated goals for presentation to the membership at the Fourteenth Session to be held in Paris in 1988.

Attachment

Mission and Goals

The mission of SIO in the area of automation must derive from the general aims and purposes outlined in its statutes. These are:

(a) To establish, maintain, and strengthen relations between archivists of international organizations and between all institutions, professional bodies and other organizations or persons concerned with the custody, organization, or administration of archives of international origin, public or private, wheresoever located;

(b) To promote all measures for the preservation, protection and defence against all hazards of the archival heritage of international organizations, and to further the progress of all aspects of the administration and techniques of the preservation of such archives by facilitating the exchange of ideas and information on problems relating to archives of international origin;

(c) To facilitate the more frequent use of such archives and their more effective and impartial study by making their contents more widely known and by encouraging greater freedom of access;

(d) To promote, organize and co-ordinate activities in the field of archival administration in international organizations, including meetings of the archivists concerned;

(e) To encourage professional training of staff responsible for the archives of international organizations;

(f) To co-operate with all concerned with the documentation of international organizations.

It is proposed that in order to best accomplish the Section's mission in the area of automation, a Five-Year Plan should be prepared to achieve the following goals:
Mission and Goals (continued)

1. To facilitate adequate automation-related training for archivists and records officers of international organizations to permit them to effectively apply modern technology in their work, and successfully carry out their duties and responsibilities with regard to machine-readable records.

2. To enable archivists and records officers of international organizations to clearly define their role relative to other information services and to achieve recognition of their proper role within their respective organizations.

3. To promote the establishment of standards to facilitate the exchange of electronic data within organizations—as well as between them—for purposes of facilitating more effective records management, and the preservation and protection of and access to electronic records.

4. To create a mechanism for the dissemination of automation-related information to archivists and records officers in international organizations (on either a selective basis or upon demand), and also for the timely and complete exchange of information between and among archivists and records officers in international organizations.

5. To adopt guidelines for the management and preservation of electronic and mixed-media records (file series in which the information exists in part in textual form and in part in machine-readable form).


Agenda Item 5.2

The Section of Archivists of International Organizations requests its Executive Board to transmit to the U.N. Administrative Committee on Coordination (ACC) the attached Report dealing with Cooperation within the United Nations System in the Field of Archives (ACC Decision 1984/15). It also requests that copies of the Report be transmitted to the Executive Heads of the 29 organizations of the U.N. system from which information was solicited pursuant to ACC Decision 1984/15.
VIII. PROPOSAL FOR A TECHNICAL PANEL ON RECORDS MANAGEMENT

31. The Committee had before it a secretariat background paper (ACCIS 87/012) which pointed out that, as a result of the rapid evolution of information and communication technologies, archives and record management practices, based essentially on information on paper supports, were no longer adequate to accommodate the new technologies. It was therefore suggested that a technical panel on records management be established to consider how an orderly transition to new forms of records management could best be ensured, both in terms of securing immediate gains in efficiency and to long-term advantages including historical purposes. The Committee agreed that the problem was common to a number of organizations, that it was both important and urgent, and therefore favoured the establishment of such a technical panel.

32. The broad mandate of the technical panel would be to study the impact of the new technologies on established records management methods and services and ways and means of effecting a smooth and efficient transition from a primarily paper-supported environment to one increasingly oriented to electronic communications and records. At the same time, the Committee recognized that a number of aspects of the problem were closely related to the activities of TP/COM and that duplication of activities and ensuing confusion should be avoided. To this end, the Committee agreed that the panel on records management should proceed to a preliminary identification of issues to be considered and frame a provisional work plan before the end of 1987. For its part, TP/COM should likewise prepare a plan of work regarding issues of records management arising from its activities relating to computer-assisted communications. Both work plans should be submitted to the ACCIS Steering Committee which would on that basis ensure a proper interfacing between the related activities of the two panels, while avoiding duplication or conflicts of competence.

33. The Committee agreed that the new panel would be established with the World Bank as the lead agency. The core members of the Panel would include: UN, UNDP, UNESCO, ILO, IMF, WHO and ICC, ECLAC, UNDRO, ITC and IFAD would participate in the new panel as associate members.