Inventorying Magnetic-Media Records

By EVERETT ALLDREDGE

O VER THE YEARS the archival profession has developed a degree of unanimity on what constitutes a satisfactory minimal inventory of records holdings in an archival repository. Over a much shorter period, records management specialists have established a doctrine on what constitutes an adequate minimal inventory of records in the creating and maintaining establishment. In both cases the point of view has been largely confined to textual records. Archivists with large holdings of maps, charts, engineering drawings, sound recordings, still pictures, and motion pictures have had to develop a different methodology for inventorying these types of records, and there is reason to believe that this methodology, too, will soon become standard.

Before 1968, no body of theory existed on what constitutes an adequate inventory of magnetic tapes, although they are an indispensable ingredient of computer-produced information. In 1968 the National Archives and Records Service thought it time to begin work on an inventory theory and began a project to find out what information existed in machine language form in the Federal Government, and where. The project carried with it the inevitable corollary considerations of whether any of the information was archival in nature and, if so, what the National Archives should do about it. The writer was the NARS project leader, and this article comes from firsthand experience, albeit the experience was experimental and the NARS project still remains a pilot adventure. The latter phases of this project, although consequential to the reader, would make a long story in itself. This article will, therefore, concentrate on the questions one must ask and have answered in order to put together an inventory of magnetic-media records.

The author, a Fellow and past President of the Society, retired in 1971 from the National Archives and Records Service as Assistant Archivist for Records Management, after 31 years of Government service. Author of many articles in management and paperwork management journals, he will be remembered by *American Archivist* readers for his "Combining Archival and Records Management Terminology" and his thoughtful presidential address to the Society, published as "Still To Be Done."

Computer tapes are usually kept in see-through, plastic reel boxes in the creating and maintaining organization by a "tape librarian" in a "tape library." This nomenclature is somewhat misleading because conventional library techniques are almost never used in computer tape custodial areas.

Tape librarians usually rely on (1) reel labels and (2) reel lists to find tapes wanted by the data processing staff. NARS has collected over 400 reel container labels, and no two are exactly alike. Generally, though, they contain a reel number and a "short title" (such as "January 1970 payroll" or "General Ledger, June–July 1970"). As the reel lists are usually derived from the reel box labels, in numerical order, the reel lists differ as widely as the labels. The lists are often cards of different sizes; they are frequently handwritten, and some tape librarians resist putting the lists into typewritten form. On the other hand, some of the big installations have computer printouts to use as reel lists.

The first requirement was to get an idea of how many tape libraries there were in the Federal Government (almost 4,000), where they were located, and what information, stated in general terms, they contained. The tape libraries held from as few as 200 reels to as many as 60,000. We in NARS thought that if these reels could be grouped by the archival "records series" concept, then a practical descriptive handle would be forged. The concept of records series turned out to be not too difficult for most of the tape librarians to master. Some tape libraries contained as few as 32 series; others more than 2,400.

To make the desired inventory manipulable, a form was designed for the tape librarians to complete, series by series. The form bears the identification number GSA Form 7036 (see figure 1). It asks 23 questions about each series. As most inventory forms for textual records contain between 7 and 10 questions, one may reasonably conclude that a magnetic tape record inventory is at least more lengthy, if not more complex.

Actually, no tape librarian had enough knowledge about the tapes in his or her custody to fill out the forms unaided. This is why the answers to inventory items 5A and 5B became so important. By learning the name of the "sponsor" (the head of the organizational unit for which the information in the magnetic tapes was designed) or the "EDP coordinator" (the principal person who designed the way the information was systematized within the magnetic tapes), the tape librarian learned who could answer the questions he or she could not.

The heart of the inventory is question 8. Broken down into 5 parts, it supplies the description of information content in the tape

series. The answers to questions 12 and 13 provide information on whether the tape series is a one-time collection of data or whether the initial data series is constantly being updated by newer and later information being fed into the tapes. The 1970 population census is an example of a one-time collection of data, while the cost-of-living index prepared by the Bureau of Labor Statistics is an example of a frequently updated data collection.

Question 9 requires that the source document(s) used be attached to the inventory form. Computer tapes are simply information transposed from textual form to machine-readable and machine-processable form. The tapes contain only that information transposed from what are now generally called "source documents," which are generally forms. For a payroll the source document is a time and attendance record, and for social security it is an initial entry of a few basic facts about an individual plus a quarterly update of earnings. The different types of source documents are rarely more than two in number; therefore to attach them to the inventory form as required by question 9 is no great burden.

A copy of a printout from the magnetic tape, the printout being a textual document, often a report, is needed for question 10. It will show how the information on the source documents has been summarized, aggregated, or otherwise formatted. Again, it is no chore to those filling in the inventory form to provide one printout (there may be several from which to choose).

Question 14 reminds the tape librarian that information taken from the source documents is usually coded—indeed, when keypunching is involved, the keypuncher may work from the coded sheet rather than from the source document itself. This question requires that a copy of this "specification" be attached to the inventory form, for the "spec" indicates what "data elements" constitute the information base of the magnetic tapes.

The needed information about the brand of computer equipment used in processing the tapes is provided by question 16, while the "BPI" in question 19 is the number of "bits per inch," which is now usually 1,600, a fairly high compaction. Questions 21 and 22 are intended to give information on how long the tapes are being kept by the creating organization. NARS received such unusual answers to these two questions that the records management staff had to make many visits and phone calls to straighten out what was admittedly an untenable records maintenance situation. Most tapes had no specified retention period, so tapes were being erased by programers who needed more blank tape, by tape librarians when the tapes had not been used for "several" months, and so on.

Depending on how Question 22 is answered, the magnetic tape

GENERAL SERVICES ADMINISTRATION
NATIONAL ARCHIVES AND RECORDS SERVICE

MAGNETIC TAPE RECORD INVENTORY

1. RECORD GROUP NO.	2. CONTROL NO. *(FOR NARS USE)*

3. AGENCY/OFFICE CREATING RECORD

4. LOCATION OF EDP INSTALLATION

5. OFFICIALS RESPONSIBLE FOR SYSTEM *(Name)*

6. BUILDING ROOM NUMBERS	7 TELEPHONE NUMBERS
A.	A.
B.	B.

A. SPONSOR

B. EDP COORDINATOR

8. DESCRIPTION OF RECORD CONTENT

A. SYSTEMS TITLE

B. FILE(S) TITLE

C. PURPOSE OF COLLECTING DATA

D. SCOPE *(Content and coverage)*

E. ARRANGEMENT-SORTING SEQUENCE *(Logical record key)*

9. SOURCE DOCUMENT(S) USED AS INPUT (Attach samples)

10. USE OF FILE OUTPUT-PUBLICATION TITLE (Attach copy)

11. DATES OF FILE

A.. FROM

B. TO

12. ONETIME STUDY OR SURVEY (S)

13. PERIODIC UPDATE (Specify period)

14. FILE SPECIFICATION DESCRIPTION (Enclose record format and data elements description)

15. DUPLICATION ELSEWHERE (Physical or content)

16A. CPU MFG.

16B. MODEL NO.

16C. TAPE DRIVE

16D. MODEL NO.

17. NO. OF REELS

18. REEL LENGTH

19. BPI USED

20. NO. OF TRACKS

21. CURRENT RETENTION PERIOD

22. RECOMMENDED RETENTION PERIOD

23. DATE REELS WILL BE TRANSFERRED TO NARS

24. USE MADE OF TAPE RETAINED OVER TWO YEARS

25. DEFENSE CLASSIFICATION AND/OR RESTRICTIONS ON USE OF TAPE

GSA FORM 7036 (REV. 3-71)

Figure 1

GSA Form 7036, Magnetic Tape Record Inventory

DATA ARCHIVES INVENTORY
(Read Instructions on reverse)

FOR NARS USE ONLY

SECTION I - PHYSICAL CHARACTERISTICS

1. TYPE OF REEL	2. LENGTH *(Feet)*	3. WIDTH *(Inches)*

SECTION II - RECORDING MODE INFORMATION

1. NUMBER OF TRACKS

2. DENSITY IN BPI

3. CHARACTER CODE USED

4. FRAME PARITY

5. MIXED MODE READING CONTROL

6. LOGICAL RECORD LENGTH

LENGTH
☐ WORDS   ☐ CHARACTERS
☐ FIXED   ☐ VARIABLE

7. BLOCKING FACTOR *(Logical records per physical block)*

8. LOGICAL AND PHYSICAL RECORD LENGTH CONTROL

	(1) NONE	(2) COUNTER	(3) SPECIAL CHARACTER	(4) DESCRIBE
A. LOGICAL	☐	☐	☐	
B. PHYSICAL	☐	☐	☐	

☐ FRAMES
☐ UNDEFINED

9A. END OF FILE PADDING CHARACTER _____ IN POSITIONS _____ TO _____ OF FIRST LOGICAL RECORD

AFTER LAST LOGICAL RECORD OR NONE _____ .

9B. IS LAST BLOCK FULL LENGTH _____ OR SHORT _____ ?

9C. INDICATE END OF REEL PADDING FIELD IF DIFFERENT FROM ABOVE

SECTION III - RECORDING SYSTEM INFORMATION

1. CPU MANUFACTURER

2. MODEL NO.

3. TAPE DRIVE MANUFACTURER

4. MODEL NO.

5. OPERATING SYSTEM USED TO GENERATE FILE

A. NAME

B. VERSION NUMBER

C. LEVEL OR TYPE

6. FILE STRUCTURE INFORMATION (See reverse for Instructions)

TM ———— (LABEL LENGTH) ———— TM ———— DATA RECORDS ———— TM ———— EOR ———— TM ———— EOF ———— TM ————

———— LABELS ————

7. FILE ID LOCATION

8. FILE ID CONTENTS

9. REEL SEQUENCE NUMBER LOCATION

10. CHECKPOINT LOCATION (If applicable)

11. LOGICAL RECORD COUNT LOCATION

12. PHYSICAL RECORD COUNT LOCATION

13. HASH OR CONTROL TOTALS (If applicable, describe)

SECTION IV - REMARKS

GENERAL SERVICES ADMINISTRATION

GSA FORM 7091 (4-71)

Figure 2

GSA FORM 7091, DATA ARCHIVES INVENTORY

record inventory goes into a new phase.   If there is reason to believe that the information on any tapes (i.e., records series) is "archival," then the inventory—for that particular series—goes into greater depth.   GSA Form 7091 (see figure 2) was designed to provide the needed Data Archives Inventory.   The number of such tentative "archival" tapes in the Federal inventory is comparatively small. The Magnetic Tape Record Inventory eventually covered about 6 million reels of tape; the Data Archives Inventory covers only 77,000.   All 77,000 reels eventually will be offered to the National Archives for accessioning.   At the time of offer the Records Appraisal Division of the National Archives will decide whether the tapes are of permanent value and, therefore, should be accessioned. It is safe to predict that the appraisers will decline many of the offers.

The Data Archives Inventory form contains the following instructions for its completion:

SEC. I.   PHYSICAL CHARACTERISTICS

1. Type of reel	Cassette, cartridge, reel, etc.

SEC. II.   RECORDING MODE INFORMATION

1. Number of tracks	7, 9, other
2. Density in BPI	128, 200, 250, 556, 800, 1600, other
3. Character code used	IBM, BCD, FIELDATA, BINARY, EBCDIC
4. Frame parity	even, odd, mixed
5. Mixed mode reading control	If parity varies from block to block, describe how software determines the parity of the block to be read. Example is lookahead bits in the IBM 7090 series.
7. Blocking factor	If fixed number of logical records per block, show number. If variable number of records per block, describe control technique in 8 below or use additional sheet if necessary.
8. Logical and physical record length control	If blocks and records are fixed length with no software control characters, mark "NONE". If blocks or logical records have counter fields specify mode with reference to system software manual; e.g., IBM 7070 IOCS, type 4 records.

SEC. III.   RECORDING SYSTEM INFORMATION

5. Operating system used	Name DOS, MOD I Extended, etc. Version number Revision 16 Level or type PCP, HASP, MFT-II, MVT
6. File Structure Information	Most software systems allow considerable variation in label content and placement even when using "standard" labels. The following example shows a typical file structure and how it should be depicted on the reverse side. IHDR (84 characters) TAPE MARK CHECKPOINT/RERUN TAPE MARK DATA RECORDS TAPE MARK EOF (end of file label) TAPE MARK TAPE MARK Such a sequence of records and tape marks should be shown on line 6 as follows: TM Ø LABELS 1 (LABEL LENGTH 84) TM 2 DATA RECORDS TM 1 FOR TM___ EOF TM ___
10. Checkpoint location	Show, for example, as "between the two tape marks immediately preceding the data records."
13. Control totals	Indicate whether and where the record counts are found in the trailer labels or within a logical record such as the one containing the sentinal characters. Hash and control totals other than record counts are usually non-standard in all software systems. Therefore, indicate what data fields are used and indicate whether binary or decimal arithmetic is used.

The significance of the information provided by the Data Archives Inventory is that tapes, if and when accessioned by the archival institution, can be used by persons not in the creating organization.

In the creating organization a sizable volume of operational documentation is produced to permit effective use of the tapes and updating and alteration of information content. For archivists who need to know more about the documentation produced in an automatic data processing unit, a book by Peter Zuckerman, *System Life Cycle Standards: Forms Method,* (Brandon/Systems Press, Inc., Princeton, N. J., 1970) is recommended.

The archival institution needs only a *part* of the documentation produced by the creating organization in order to use the magnetic tapes created, and the Data Archives Inventory points up the needed part. It does this through a section devoted to recording mode information and a section on recording-system information. These sections are not as formidable as they might at first appear, as can be discovered by taking the form to a computing facility and watching systems analysts there complete it for any records series suggested.

It is already the experience of the National Archives that the most common type of reference service is that of requesters wanting to purchase (1) entire reels, or (2) parts of reels. Most users, therefore, have access to computers. Accompanying the copy of the reel from the archival institution must be a documentation booklet giving the requester the information he needs to use the tape. The Data Archives Inventory enables the archivist to prepare the documentation booklet speedily and accurately.

It is also the experience of the National Archives that the preparation of a magnetic tape inventory is not a one-time undertaking. Some of the participating organizations admittedly will not do a good job the first time around. Records managers will want a second try, if not for all the organization, then at least for parts of it. In some organizations the number of new tapes is increasing at a high rate, and the records managers there will want to update the inventory every two or three years. As people become better acquainted with what is involved, updating, on whatever basis, becomes easier to effect.

This article does not pretend to answer all of the technical questions that arise in taking an inventory of magnetic-media records. The National Archives now has a well-established Data Archives Branch with a staff prepared and willing to answer your technical questions.