DOE/TIC-4581-R4

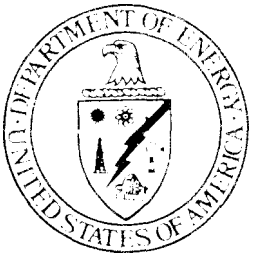
EDB

DOE/TIC--4581-R4

DE80 002729

Magnetic Tape Description

June 1980



UNITED STATES DEPARTMENT OF ENERGY
TECHNICAL INFORMATION CENTER

DOE/TIC Bibliographic Reference Manuals

A series of publications in the form of guides and manuals has been developed by the Department of Energy Technical Information Center to create the Energy Data Base and other computer-based data collections, as well as the announcement publications and bibliographies covering DOE program interests. The individual titles, listed below, are for sale from the National Technical Information Service, Springfield, VA 22161.

DOE/TIC-85-R13, *Energy Information Data Base: Report Number Codes,* September 1979, $16.00.

DOE/TIC-4579-R11, *Energy Information Data Base: Serial Titles,* in press.

DOE/TIC-4584-R4, *Energy Information Data Base: Energy Categories,* March 1980, $7.00.

DOE/TIC-4602, *Energy Information Data Base: DECOL Manual (Descriptive Cataloging On-Line),* January 1980, $14.00.

DOE/TIC-7000-R4, *Energy Information Data Base: Subject Thesaurus,* September 1979, $22.50.

DOE/TIC-4581-R4, *DOE Energy Information Data Base: Magnetic Tape Description,* June 1980, $5.00.

TID-4583-R1, *Energy Information Data Base: Guide to Abstracting and Indexing,* February 1978, $6.50.

TID-4585-R1, *Energy Information Data Base: Corporate Author Entries,* June 1978, $15.00.

Available from the National Technical Information Service, U. S. Department of Commerce, Springfield, Virginia 22161.

Price: **Paper Copy** $5.00
Microfiche $3.50

Printed in the United States of America
USDOE Technical Information Center, Oak Ridge, Tennessee

DOE/TIC-4581-R4
June 1980

DOE ENERGY INFORMATION DATA BASE
MAGNETIC TAPE DESCRIPTION

Charles W. Duncan, Jr.
Secretary

Ruth M. Davis
Assistant Secretary
Resource Applications

Joseph G. Coyne
Manager
Technical Information Center

Noboru Kawakami
Editor

About the Technical Information Center

The Technical Information Center, as the national information center for the Department of Energy, manages, publishes, and disseminates scientific, technical, and practical energy information acquired from worldwide sources. It builds and maintains energy information data bases used to prepare abstracting journals and bibliographies. The data bases are also searchable via the Department's on-line retrieval system, DOE/RECON. This information is provided to DOE program offices; contractors; scientists; engineers; educational, commercial, and industrial communities; and the general public. The Technical Information Center conducts an active announcement and marketing program. It establishes policy and administers control of a DOE-publications inventory and registry. TIC also develops and evaluates new and effective information systems and services as required by scientific and technical management within the Department of Energy.

About the DOE Energy Information Data Base: Magnetic Tape Description

In the processing of bibliographic citations, abstracts, and subject content indicators (descriptors and data base categories), information is encoded at TIC for computer storage, manipulation, and retrieval using both on- and off-line systems.

Expansion of TIC's data files to include nonnuclear information is relatively recent, beginning in 1974. TIC's selection and announcement of technical information now corresponds to DOE's research and development interests that include all areas of energy. To make the files as comprehensive as possible, both current and background information is included. As a part of its responsibility to disseminate as widely as possible technical information related to DOE's programs, TIC provides copies of magnetic tapes representing the completed bibliographic intake at TIC for a two-week period. The number of bibliographic records included in each tape increment is expected to vary between 2,000 and 7,000 citations.

Records are loaded on the update tapes and are sequentially numbered after being sorted by the initial subject category codes as described in Report DOE/TIC-4584, *Energy Information Data Base: Energy Categories.* Tape users are cautioned to review ALL category codes for an item on the tape since the sort is by initial category code only. The tape format conforms to American National Standard Z39.2-1971 with the number of citations varying for the various dissemination tapes developed. This document is also intended to be used as a guide for developing input tapes for submission of data to TIC.

Published by the **United States Department of Energy, Technical Information Center**
P. O. Box 62, Oak Ridge, Tennessee 37830

CONTENTS

DOE ENERGY INFORMATION DATA BASE: MAGNETIC TAPE DESCRIPTION

1. SCOPE AND PURPOSE

1.1 General

This document describes a format for disseminating bibliographic citations, abstracts and subject content indicators for energy information on computer readable magnetic tape issued after July 1976. The magnetic tapes contain information incorporated into the DOE data base. This base is used by TIC to generate multiple publications and to support information retrieval activities. Information, or data, pertinent to one bibliographic reference will be called an "entry." The entries are logically composed of data elements (such as title, authors, etc.). This format supersedes a previous version, used since 1974, as described in Report TID-4581-R3. The format has been revised to permit incorporation of DOE/TIC bibliographic data encompassing the broader scope of information related to the development or use of energy resources.

1.2 References

This format is intended to conform to ANSI Z39.2-1971, American National Standard for Bibliographic Information Interchange on Magnetic Tape. The coding for character representation is superimposed upon ANSI X3.4-1968, American National Standard Code for Information Interchange, but the ANSI code itself does not completely define the set of characters needed. These references to other standards are made only for background information; this document stands alone for delineating the dissemination format and code of the TIC tape. Definition of and details concerning the information being disseminated, however, will not be treated in this document, except in a general way and by reference to other sources. That is, nothing here will explain TIC cataloging and indexing rules—the user must rely on other references to understand the content and logic of this information; only its physical representation is herein described.

2. MAGNETIC TAPES USED IN DISSEMINATING DOE/TIC INFORMATION

2.1 Tapes

These tapes will be conventional magnetic tapes commercially available in the USA and are ½ inch wide, size II, 2400 feet long on 10½ inch reels, Class II nine track, 6250 CPI, with Type I wraparound tape reel band.

2.2 Recording

Recording of data on these tapes will be done by magnetic tape units generally described in the USA as *IBM compatible.* The recording scheme will be 1600 CPI, phase encoded. This scheme is in compliance with the *American National Standard Recorded Magnetic Tape for Information Interchange,* ANSI-X3.39-1973. This recording method is totally hardware controlled.

2.2.1 Blocks

A block, sometimes called a physical record, is that contiguous string of bytes between gaps used for recording data. Blocks shall not exceed 2044 bytes, inclusive of a leading 4-byte block count field with each block. It is possible that some entries will require more than one block to contain them; thus any given entry may span multiple blocks. The coding scheme for indicating these multiple block entries is described in paragraph 4.2.1. Blocks may contain more than one entry. No block will contain fewer than 28 bytes. See Figure 5 for block/entry relationships.

2.2.1.1 *Noise Blocks.* Since magnetic tape surfaces may become locally damaged, a scheme for ignoring "bad" spots should be employed. Any blocks that are fewer than 28 bytes in length may be safely ignored as noise, since no valid data will be disseminated in blocks shorter than this. Parity and/or redundancy checking schemes may fail and should be waived if a noise block is encountered.

2.2.1.2 *Block Count.* A 4-byte block count will be at the beginning of each block. This will be a decimal number, right adjusted, and will not exceed 2044 or be less than 28 bytes. This count includes its own 4-byte field.

3. CODES

3.1 General

Each frame recorded on the magnetic tape represents a character or byte. These characters are represented by 8 bits (or an 8-level code) which may be arranged in 256 unique combinations of 0 or 1. Of these 256 possible character codes, 120 codes are used to indicate *printing characters.* These are the characters used in printing and are shown in Table 1. Other characters are used as "functions" or controls, and these are also tabulated in Table 1. All characters used, whether as printing characters or control characters, are tabulated along with their code arrangement in Table 2. These codes have been superimposed upon the ANSI X3.4-1968 code for information interchange (ASCII code). That is, where characters and associated codes are defined in the ANSI standard they are used; additional characters not covered by the standard but required for printouts have been assigned other codes, as delineated in Table 2. To accommodate users who prefer a 7-bit code structure, all non 7-bit ASCII characters will be preceded by an Escape (ESC) Code and a Font Identifier. The Font Identifier will be binary one (00000001). This scheme provides for future character set expansion. Effectually, bit 8 can be ignored if desired.

Table 1 CHARACTER SET

Printing Characters

A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

a b c d e f g h i j k l m n o p q r s t u v w x y z

1 2 3 4 5 6 7 8 9 0 . : ; , / ♀ = + – → $ ♂ % ∫ √ *

> < [ ] ( ) _ (underline) ' (quote, apostrophe)

α β Δ Σ π Ξ ω γ ν Λ Ω μ

superscript: ¹ ² ³ ⁴ ⁵ ⁶ ⁷ ⁸ ⁹ ⁰ ⁻ ⁺

subscript: ₁ ₂ ₃ ₄ ₅ ₆ ₇ ₈ ₉ ₀

Function Characters

ESC	Escape or Shift Code
SP	Space
FT	Field Terminator
RT	Record Terminator
TM	Tape Mark
US	Unit Separator

Note: In general, the Unit Separator (US) will be used as a separator for words or numbers within a data element, e.g.; to separate descriptors.

Table 2 CHARACTER CODES

BITS	8	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
	7	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
	6	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
	5	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1
4 3 2 1	Col / Row	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0 0 0 0	0			SP	0		P		p							0	0
0 0 0 1	1				1	A	Q	a	q							1	1
0 0 1 0	2				2	B	R	b	r							2	2
0 0 1 1	3		TM		3	C	S	c	s							3	3
0 1 0 0	4			$	4	D	T	d	t							4	4
0 1 0 1	5			%	5	E	U	e	u							5	5
0 1 1 0	6				6	F	V	f	v							6	6
0 1 1 1	7			'	7	G	W	g	w							7	7
1 0 0 0	8			(	8	H	X	h	x							8	8
1 0 0 1	9			)	9	I	Y	i	y							9	9
1 0 1 0	10			*	:	J	Z	j	z				α			+	Σ
1 0 1 1	11		ESC	+	;	K	[	k					β			-	μ
1 1 0 0	12			,	<	L		l					γ		♂	√	ν
1 1 0 1	13		RT	–	=	M	]	m							♀	Δ	ω
1 1 1 0	14		FT	.	>	N		n							→	Λ	π
1 1 1 1	15		US	/		O	_	o							∫	Ω	Ξ
		← 7 Bit ANSI Standard →								← TIC Font 001 Extension →							

NOTE: An example of the character string to represent the text string $\pi^0 \rightarrow 2\gamma$ on tape would be ESC ① πESC ① ⁰ ESC ① → 2ESC ① γ. ① is the Font Identifier One (Bits 00000001).

4. FORMAT

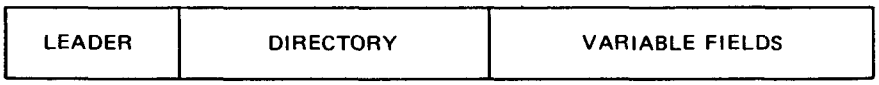
4.1 General

Each entry will be made up of three basic, logical sections:

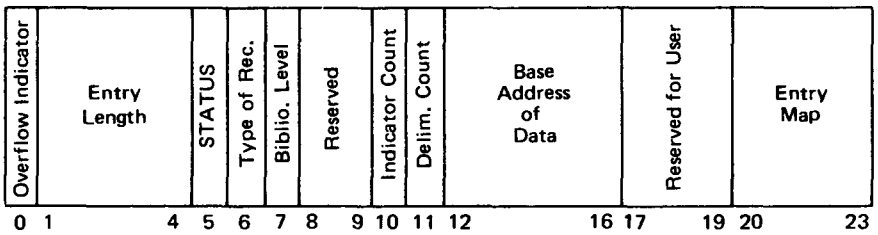
1. Leader (fixed length)
2. Directory (a variable number of finders)
3. Variable Field (where the information itself is placed)

The leader contains fields useful in processing the remainder of the entry and contains little information except for the type and bibliographic level of the entry. The directory may be considered as a series of finders required to identify and locate the actual information placed in the variable field. Schematic representations are shown in Figure 1.

a. The interchange format is schematically represented below:



b. The leader has the schematic format shown below:



c. The directory consists of a series of finders. There must be one finder in the directory for each variable field. Each finder has the schematic format shown below:

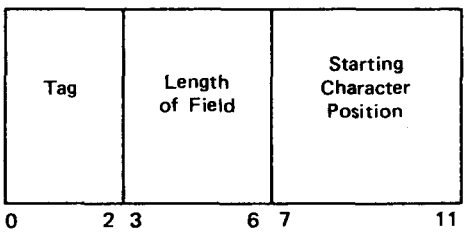


Fig. 1 Record schematic.

4.2 Leader

The general format for the leader is shown in Figure 1. The leader is 24 characters in length and contains basic control information. In case an entry spans multiple blocks, the leader appears only in the first block.

4.2.1 Overflow Indicator (0)

If this single-digit decimal number is one, then the entry is completed in that block; otherwise the entry will continue into, or span, subsequent blocks, and this indicator will be the number of blocks. For example, if three blocks are required to contain a complete entry, the overflow indicator will be "3."

4.2.2 Entry Length (1-4)

This 4-digit decimal number equals the length of the entire entry including itself and will not exceed 9999. It does not include the Longitudinal Redundancy Check or Cyclic Redundancy Check bytes associated with each block or the block count word at the beginning of each block.

4.2.3 Status (5)

This single character field will always be set to "N" (for New), "C" (for Change) or "D" (for Delete). Status codes C and D imply that the entry has been previously disseminated and should be changed or corrected with the current data (C) or deleted entirely (D).

4.2.4 Type of Entry (6)

This single character field will indicate the literature type of the entry through the following codes:

Literature Type	Type of Record (6) Code
Report	R
Analytics of a Report	U
Journal Article	J
Journal Article Translation	V
Book	B
Analytics of a Book	Y
Patent	P
Engineering Drawings	E
Dissertation	D
General Translation	T
Analytics of a Translation	Z

4.2.5 Bibliographic Level (7)

This field will contain a code denoting the hierarchically lowest (or only) bibliographic level assigned by the cataloger. The following code will be used.

Bibliographic Level	Code (7)
Analytic	A
Series	S
Monograph	M
Collective	C

4.2.6 Reserved for Future Use (8-9)

These characters will be spaces.

4.2.7 Indicator Count (10)

This position will contain the digit "0" (zero). This feature of the ANSI Z39.2 standard is not used.

4.2.8 Delimiter Count (11)

This position will contain the digit "0" (zero). This feature of the ANSI Z39.2 standard is not used.

4.2.9 Base Address of Data (12-16)

This is a 5-digit decimal number which equals the combined length of the leader and the directory. This number "points to" or locates the start of the variable fields containing the data. For example, an entry with 10 variable fields would have a base address of 145.

4.2.10 Reserved for Users (17-19)

These characters will be spaces.

4.2.11 Entry Map (20-23)

This field is provided by the ANSI standard to allow a variable directory entry scheme. All directory entries (finders) for this data will be fixed. The entry map will always equal "4500." The "4" indicates a 4-digit "length of field" portion and "5" indicates a 5-digit starting character position portion of each finder (see Figure 1 and Sec. 4.3.1).

4.3 Directory

The directory consists of a series of fixed fields, called *finders* (or entries by ANSI). There will be one and only one finder for each subsequent variable field (data element) in the entry.

4.3.1 Finders

Each finder consists of 12 characters. (See Figure 1 for a schematic.) Each finder will consist of a tag, length of field, and starting character position, in that sequence. Finders point to, or "find," data elements in the variable field. In this implementation finders will be in the same order as their corresponding data elements.

4.3.1.1 *Tag.* The tag is a 3-digit number symbolically identifying the data element recorded in the variable field. Table 3 defines the relationship between the symbolic tag numbers and the corresponding data elements.

4.3.1.2 *Length of Field.* Length of field is a 4-digit decimal number equal to the length (in bytes) of the data element in the variable field to which it corresponds (including the element's field terminator).

4.3.1.3 *Starting Character Position.* This field is a 5-digit decimal number equal to the position of the first character in the corresponding variable field data element. This position is relative to the base address (position 12-16 of the leader).

Example: The starting character position of the first data element for an entry (tag 001) is 00000. Note that the starting position is beyond the leader and directory.

4.3.2 Field Terminator

A field terminator character terminates the last entry in the directory.

4.4 Variable Field

Following the leader and directory, the entry is presented in a variable field made up of one or more data elements. (The directory itself is technically a variable field but is of different construction.)

Table 3 TAG NUMBER AND DATA ELEMENT TITLE

Tag No.	Data Element Title	Tag No.	Data Element Title
001	EDB File Number	390	Pages
010	TIC Item Identification Number	410	Translation Note
011	EDB Tape Issue Number	420	Language
020	Type of Item	430	Availability, Price
030	Classification	440	Drop Notes
040	Literary Indicator	450	Conference Title
060	Personal Author (A)	460	Conference Place
070	Personal Author (M)	470	Conference Date
090	Primary Title (A)	490	Thesis Statement
100	Subtitle (A)	500	Project ID Code
110	Primary Title (M)	510	Distribution
120	Subtitle (M)	520	Report Origin
130	Primary Title (S)	530	File (Selected for)
150	Primary Report Number	540	EDB Category Number
170	Affiliation (A)	550	Source of Bibliographic Information Code
190	Affiliation (M)	560	Country of Affiliation Code
200	Original Title (M)	570	Country of Publication Code
210	Secondary Report Number	590	INIS TRN Number
220	Patent Number	600	INIS Type
230	International Classification Code	610	INIS Category Number
240	Contract Number	620	Title Augmentation
250	CODEN	640	NTIS Code (NTS)
260	Abbreviated Journal Title	700	Corporate Code (INIS)
300	Assignee	710	Corporate Author (INIS)
310	Place of Publication	750	TIC Utilization
320	Name of Publisher	800	Descriptor Count
340	Volume (Journal)	801	General Descriptors (see paragraph 4.4.3.3)
350	Issue (Journal)	802	Descriptor Splits (see paragraph 4.4.3.3.1)
360	Page(s) (Journal)	803	Descriptor Upposted Terms
370	Date	950	Abstracts
380	Filed Date (Patents)		

4.4.1 Structure of the Variable Field

Each element of the variable field (data element) consists of a character string (all composed of printing characters and associated function characters) followed by a single field terminator character (FT).

4.4.2 Sequence

Data elements will occur in the ascending tag sequence, i.e., the same sequences as their corresponding directory entries. See Table 3 for tag numbers and their associated data element titles.

4.4.3 Variable Data Element Fields

4.4.3.1 *Control Number Field.* The control number field (which corresponds to finder tag 001) is a special field intended to be the key field for ordering the entries within a file. It uniquely identifies each entry and is constructed of year, colon, and sequence number (the order in which items are pulled or ordered for the file).

Example: 76: 000001

4.4.3.2 *Data Element Fields.* The remainder of the entry will consist of variable data elements, each consisting of a variable length string of characters terminated with an (FT) character, except that the last data element in the entry will end with a (RT) character. Not all possible tags and their associated data elements will necessarily be present in any one record. The Appendix contains data element definitions for all tag numbers listed in Table 3.

4.4.3.3 *Subject Indexing Scheme.* Descriptors tagged 801 and/or 802 may be marked, or labeled, as subject headings (M) and/or qualifiers (Q) for purposes of preparing printed subject indexes. Descriptors so labeled will be followed immediately by a colon (:) and the assigned labels. Thus, the colon and all characters following it (until the beginning of another descriptor or the field end) are *not* part of the descriptor but constitute information required in subject index preparation.

A given descriptor may be labeled as a subject heading (only one) and/or a qualifier for one or more headings. If a descriptor is used as a subject heading and no corresponding qualifier is provided, it will be followed by a colon, a space, and an M. If the subject heading is to be qualified, it will have a number along with the M to allow qualifiers to be associated with it (M1, M2, M3, etc.); this number will immediately follow the M and may be one or more decimal digits. If the descriptor is to be used as a qualifier of one or more headings, it will be labeled with Q's followed by numbers that correspond to the numbers of the subject headings it qualifies (Q1, Q2, Q3, etc.). In the following examples, note that commas are used to separate multiple labels applied to one descriptor. Note that (US) signifies the Unit Separator Character.

Although intended primarily for the preparation of printed subject indexes, the "M-" and "Q-" labeled descriptors might be of value to tape users who may wish to derive formats to display both the citation and indexing terms together. Such descriptor pairs have been selected and precoordinated by indexers, and their semantic relationships have thus been positively identified.

In preparing subject indexes for special publications, the labels A, T, S, and X are sometimes used in the same manner as the M label. In tape processing, these labels should be handled in the same manner as M's. The label D is sometimes used to denote numerical data.

4.4.3.3.1 *Examples of Subject Descriptor Indexing.*

Entry with no Splits:

Tag 800 element = 0009
Tag 801 element = ABUNDANCE(US)BACTERIA:M4(US)CHEMICAL ANALYSIS:Q2,Q3
(US)CHLOROPHYLL:M1(US)BIOCHEMICAL REACTION KINETICS:
Q1,Q4(US)BIOSYNTHESIS:Q1(US)SEAWATER:M3(US)WATER: M2

Tag 803 element = MICROORGANISMS(US)PORPHYRINS(US)PHYTOCHROMES(US)REACTION
KINETICS(US)SYNTHESIS(US)WATER(US)HYDROGEN COMPOUNDS(US)
OXYGEN COMPOUNDS

Entry with General and Specific Splits:

Tag 800 element = 0008
Tag 801 element = GASES(US)HIGH TEMPERATURE
Tag 802 element = RHODIUM OXIDES:M1(US)SPECTRA:Q1
Tag 802 element = PALLADIUM:M2(US)VAPOR PRESSURE:Q2
Tag 802 element = PALLADIUM OXIDES:M3(US)THERMODYNAMICS:Q3
Tag 803 element = FLUIDS(US)OXIDES(US)RHODIUM COMPOUNDS(US)PLATINUM
METALS(US)THERMODYNAMIC PROPERTIES(US)PALLADIUM
COMPOUNDS

4.5 Entry Separator

Each entry will terminate with a record separator character. This does not apply to label records.

4.6 Order

Entries will be ordered in *ascending* sequence by the File Number (tag 001). There is no way to identify or order blocks as independent entities, since some blocks may contain "overflowed" portions of an entry and thus have no leader or control number.

4.7 Padding

Since no block may be less than 28 characters (including the 4-byte block count) and, potentially, an entry's last data element(s) may not fill the 24-byte minimum, some provision for padding must be made. Any unfilled bytes with data elements at the end of a block will be made up of spaces. The RT character will immediately follow the last data element, and, if necessary to pad out to the minimum block length of 28 characters, spaces will be used.

5. LABELS

5.1 General Organization

Each dissemination tape will be identified by label records recorded on the tape. Figure 2 outlines the overall scheme. There are two types of labels:

a. Volume labels, which identify the physical tape.

b. File labels, which identify the file, i.e., the information recorded on the tape.

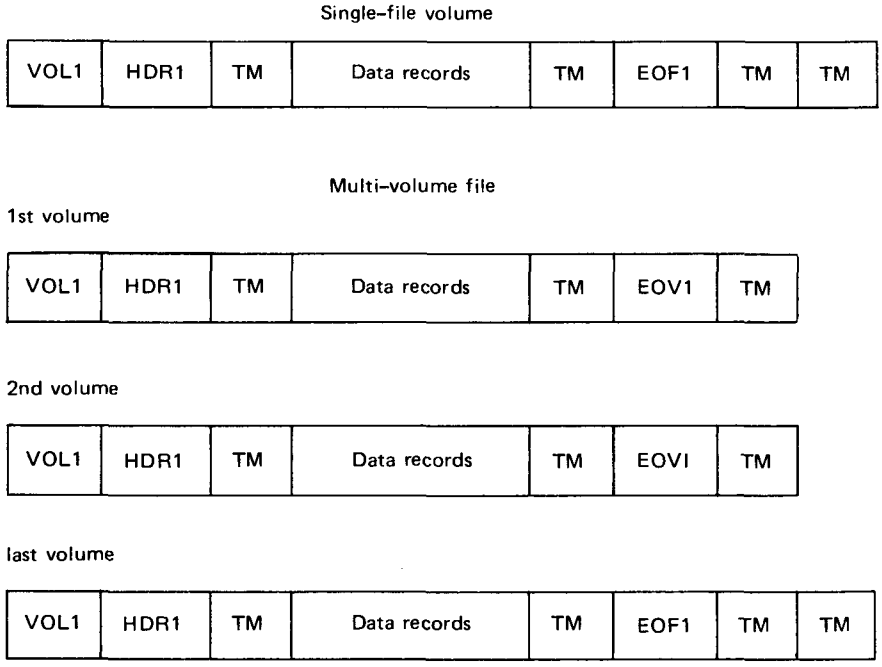
Depending on their position relative to the actual information, labels are further defined to be *header labels* if they precede the data, and *trailer labels* if they follow the data.

A file occupying more than one reel of tape is referred to as a "multi-volume" file.

This standard will cover both single file volumes (where all the data will fit onto one tape reel) and multi-volume files (where multiple reels are required). Although the usual case will be single-file volumes for each tape issue, cumulative entries may require multi-volume files. Schematic representations of these file arrangements are shown in Figure 2.

The following abbreviations are used:

VOL1: volume header label
HDR1: file header label
EOF1: end-of-file trailer label
EOV1: end-of-volume trailer label
TM: *tape mark*



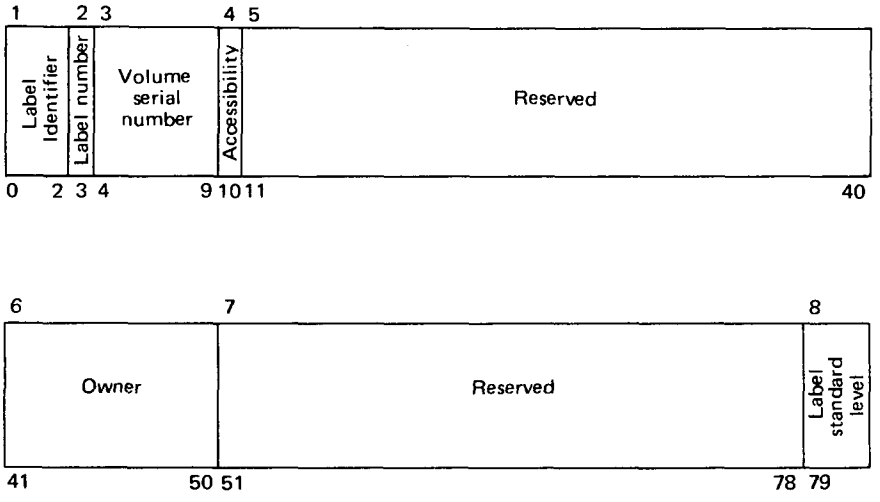
Note: The VOL1 label must be the first data recorded on the tape, i.e. it must follow immediately the reflective spot marking the physical beginning of the tape.

Fig. 2 Schematic of tape labels and file organization.

5.2 Label Format and Content

5.2.1 Volume Header Labels (VOL1)

The volume header label (VOL1) is illustrated in Figure 3. No general standards will be dictated for field 3, volume serial number, or field 6, owner; however, conventions may be standardized for particular users as this standard is implemented.



Field	Name	Length	Description
1	Label identifier	3	Will be VOL
2	Label number	1	Will be 1 (one)
3	Volume serial number	6	Six alphanumeric characters
4	Accessibility	1	Will be space, i.e., unlimited access
5	Reserved	30	Reserved for operating systems use
6	Owner	10	Ten alphanumeric characters
7	Reserved	28	Reserved; will be spaces
8	Label standard level	1	Will be 1 (one)

Fig. 3 Volume Header Label (VOL1) details.

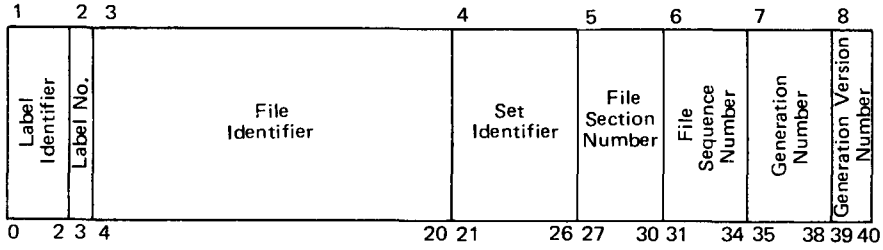
5.2.2 File Header Label (HDR1)

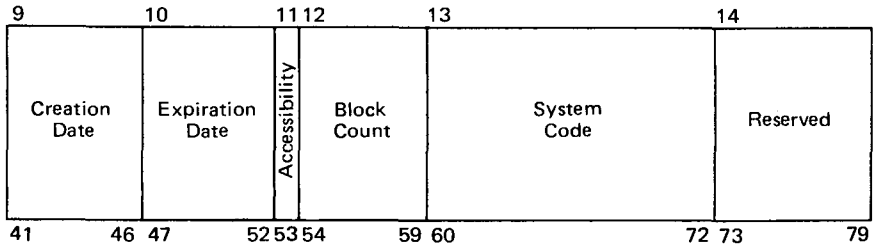
The file header label (HDR1) is defined by Figure 4.

5.2.3 End-of-File Trailer Label (EOF1)

It has the same format as the File Header Label (see Section 5.2.2).

Field	Name	Length	Description
1	Label identifier	3	Will be EOF
2	Label number	1	Will be 1 (one)
3-11	(Same as corresponding fields of File Header Label)		
12	Block count	6	A decimal number equal to the number of data blocks (exclusive of labels and tape marks) since the last HDR label recorded for the file
13-14	(Same as corresponding fields of File Header Label)		





Field	Name	Length	Description
1	Label identifier	3	Will be HDR
2	Label number	1	Will be 1 (one)
3	File identifier	17	Not used
4	Set identifier	6	This field must contain the volume serial number of the first or only volume of the file. In a multi-volume file this field is the same in every volume.
5	File section number	4	A decimal number (starting with 0001 for the first or only volume containing the file) which is increased by 1 on each subsequent volume of the file.
6	File Sequence number	4	A decimal number indicating the file sequence
7	Generation number	4	These fields are not used.
8	Generation version number	2	These fields are not used.
9	Creation date	6	A space followed by 2 digits indicating the year, followed by 3 digits (001-366) indicating the day within the year the file is created (e.g., 70001 = 1 Jan 1970)
10	Expiration date	6	A space followed by 99366
11	Accessibility	1	Will be space
12	Block count	6	Will be zero (000000)
13	System code	13	Not used
14	Reserved	7	Reserved for future standardization; must be spaces

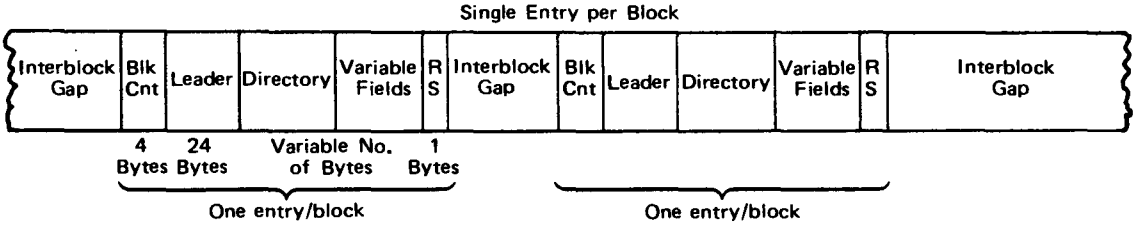
Fig. 4 File Header Label (HDR1).

5.2.4 End-of-Volume Trailer Label (EOV1)

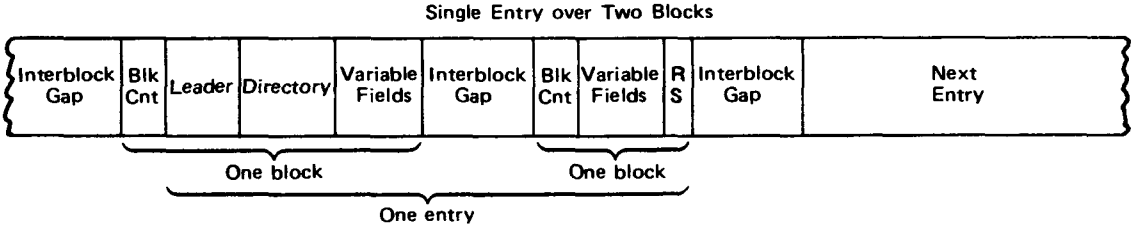
It has the same format as the File Header Label (see Section 5.2.2).

Field	Name	Length	Description
1	Label identifier	3	Will be EOV
2	Label number	1	Will be 1 (one)
3-14	(Same as fields 3-14 of End-of-File Trailer Label)		

Block/Entry Relationship



Note: RS is the record Separator Character. The Blk Cnt (Block Count) is the number of bytes in the record including the block count field of four bytes.



Note: Byte 0 of the leader in the first block will contain a 2 to indicate that two blocks are needed for the entry.

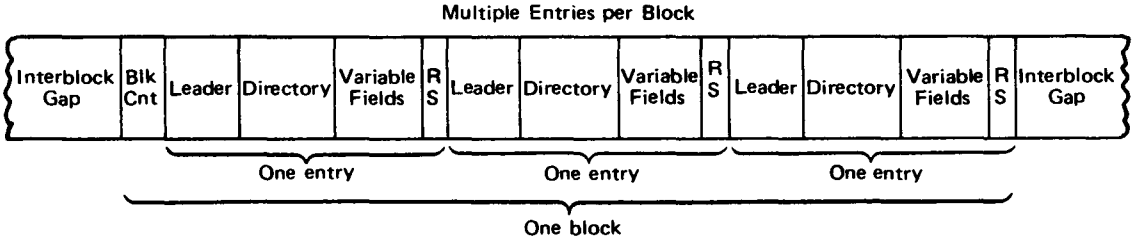


Fig. 5 Block/Entry Relationship.

APPENDIX A

DATA ELEMENT DEFINITIONS

001	File Number	The file number consists of a two-digit calendar year or volume number, a colon, and a six-digit sequence number for the files. This constitutes the ordered arrangement of the tape.
010	TIC Item Identification Number	A ten-digit number, the first two digits of which represent the year, followed by two digits for the month, and a six-digit serial number. This number is used within TIC to enable input operations to identify manuscript by the file number under which it was placed in the system.
011	EDB Tape Issue Number	A two-digit number that specifies the tape issue number.
020	Type of Item	Either one or two alphabetic characters which specify the literature type. The type of entry may be any one of the following: R = Report; RA = Analytic of a Report; J = Journal; TJ = Translation of a Journal; B = Book; BA = Analytic of a Book; P = Patents; E = Engineering Material; D = Dissertation; TG = Translation other than a monograph Journal entry; TA = Analytic of a Translation.
030	Classification	Will contain "Uncl" for unclassified entries. Although TIC has other classifications, only unclassified items will be sent to our SDI users.
040	Literary Indicator	The literary indicator may be one or more of eight different single alphabetic characters. These are: K, L, N, U, W, Z, Y, E. Multiple literary indicators are not separated.
060	Personal Author (A)	For an analytically treated item,* the personal author is listed with the last name first, a comma, space, first initial, a period, and other initials with a period after each. All additional authors are separated from the previous authors by a unit separator. After the last initial and period for an author, an affiliation enclosed in parentheses may be entered for each author. This delimiter is in free form.
070	Personal Author (M)	Identified as the personal author of a monograph, it contains the same information as in delimiter 060 above.
090	Primary Title (A)	Contains the primary title of an analytic. This title appears in upper and lower case, and the leading article, such as "A" or "The" if present in the title, has been deleted. This delimiter is in free form.
100	Subtitle (A)	Normally, if a subtitle is present for an analytic, it will be entered in this delimiter. At the present time, this delimiter is not used.
110	Primary Title (M)	The primary title of a monograph, it contains the same information as delimiter 090 above.
120	Subtitle (M)	The subtitle for monographs; it is not being entered at the present time.
130	Primary Title (S)	The primary title for a series. This delimiter is in free form.
150	Primary Report Number	The number recorded as the primary report number. The prefix is checked against the report number prefix authority. The suffix to the report number will be

		separated from the prefix by two dashes representing an "en" dash. The suffix is in free form.
170	Affiliation (A)	The affiliation (organization and/or address) of the authors for analytics. Entered in free form.
190	Affiliation (M)	The same as 170 above, but it is entered for monographs. Free form entry.
200	Original Title (M)	The original title is entered in initial capital and lower case; contains the original title for a monograph. Free form entry.
210	Secondary Report Number	As with the primary report number, the prefix of the secondary report number is entered in accordance with the report number prefix authority; it is separated from the suffix by two dashes representing an "en" dash. There may be multiple secondary reports and these will be separated by a unit separator. The suffix is in free form.
220	Patent Number	Contains two alphabetic characters for the patent office code, a space, followed by the complete patent number of up to ten characters (numerics, commas, and periods are allowable). The two alphabetic characters for the patent office are checked against a patent office table for validity.
230	International Classification Code	Entered in free form.
240	Contract Number	Contains the word Contract and a space followed by a free form input field for the contract number.
250	CODEN	CODEN entries will consist of five alphabetic upper case characters checked against a CODEN authority.
260	Abbreviated Journal Title	Inserted in the data from the CODEN or Journal authority and contains the abbreviated title of the journal.
300	Assignee	Entered in free form; multiple assignees will be separated by a unit separator.
310	Place of Publication	Contains the name of the city, a blank, followed by the standard state abbreviation. If more information is desired, it will be entered with a comma, space, and the addition of data in free form.
320	Name of Publisher	Contains the name of the publisher in free form.
340	Volume (Journal)	A free form entry but will only contain a number to indicate the volume number of the journal.
350	Issue (Journal)	Same delimiter as 340, except this is the issue number for the journal.
360	Pages (Journal)	Entered in free form and will usually contain a number or a range of numbers separated by a dash.
370	Date	Entered in the form of the day, space, three alphabetic characters of the month or season, space, and a 4-digit year. It may contain a range of months or seasons with the range indicated with a dash between the elements.
380	Filed Date (Patents)	Contains either the priority date "P" or the filed data "F". The data will be entered either with the P or F in upper case followed by a space and the date in the form as delimiter 370 above. Following the date

		there may be a space and a two character patent office code. The patent office code may not be present.
390	Pages	For books or reports, this delimiter contains the number of pages followed by a "p". For analytics of books or reports, two lower case letters "pp" and a space followed by a range of pages.
410	Translation Note	Entered in free form. Most of the entries for the translation note will be retrieved from the table but should be taken as free form by the user.
420	Language	Contains the word "In" followed by the appropriate language in language table.
430	Availability, Price	Normally obtained from a table, but free form information may also be entered by the cataloger.
440	Drop Notes	Miscellaneous information entered in free form.
450	Conference Title	Obtained from an authority but should be considered as free form entry by the user.
460	Conference Place	Obtained from an authority and should be considered as free form entry.
470	Conference Date	Entered in standard date format as data element 370 above.
490	Thesis Statement	Entered in free form.
500	Project ID Code	A two- or three-character code for the project classification.
510	Distribution	Obtained from a table and shows the standard distribution codes. Should be handled as free form entry by the user.
520	Report Origin	Contains either an upper case P for project or upper case NP for non-project.
530	File (Selected For)	Contains a three character code which identifies the various publications or purposes for which this item was selected. If it has been selected for more than one publication or purpose, additional codes will be added. Codes will be separated from the previous one by a unit separator. These codes may be EDB for Energy Data Base, PRD for Power Reactor Dockets Information, EPA for Energy Abstracts for Policy Analysis, ERA for Energy Research Abstracts, or INS for International Nuclear Information System.
540	Selected for Category	Contains a three character publication code as in data element 530 above and a six digit category number. These categories are found in DOE/TIC-4584, *Energy Information Data Base: Energy Categories.* Where there are multiple categories each will be separated from the previous category number by a unit separator (US).
550	Source of Bibliographic Information Code	Contains the code for the source which produced the bibliographic information. These data are checked against a table and contain two to four upper case characters.

560	Country of Affiliation Code	Contains two alphabetic characters to indicate the country of affiliation. These characters will be checked against country code table.
570	Country of Publication Code	Same as data element 560, except that it will be the code of the publishing country.
590	INIS TRN Numbers	This is the temporary record number which is used on items that are submitted to the International Nuclear Information System (INIS). This is placed in the record since the SDI tape is used for other purposes beyond the Selective Dissemination of Information. It will consist of US, the two-digit calendar year, and a five-digit sequence number. Example: US7600018. This will differ from the INIS number found in Data Element 750 TIC Utilization in that a zero will be added to the front of the sequence number in Data Element 750.
600	INIS Type	Contains a single character to define the type of entry. The types are B for book, C for collection or an assemblage of various documents gathered together and cataloged as a single unit, D for engineering drawing, F for film or slide (TIC does not use this type code), G for map or atlas, H for phono record or sound recording, J for journal article, P for patent or patent application, R for report, and T for a computer medium or machine readable information.
610	INIS Category Number	A three character field which consists of an alphabetic character and a two digit number. This is the subject category number for items submitted to INIS. There may be multiple categories separated by a unit separator.
620	Title Augmentation	Entered in free form and is designed to augment the title to provide more information to the user.
640	NTIS Code	A three-digit code "NTS" for items selected for NTIS.
700	Corporate Code (INIS)	Entered as a 7-digit code that is checked against the INIS Corporate Authority for validity. Multiple corporate codes are allowed separated by a unit separator.
710	Corporate Author (INIS)	Obtained from the INIS Corporate Authority by means of corporate code above. The user should consider this as being free form. Both the corporate code and the corporate author are obtained from the INIS Corporate Authority.
750	TIC Utilization	A three-letter alphabetic identifier and a dash followed by a two-digit year or volume number and a six-digit abstract number for publications. If this particular item has been published in more than one publication or included in one or more files, there may be multiple numbers. Each number will be separated from the previous number by a unit separator and identifies publications and files in which an entry is placed prior to organization of an issue tape. Example: ERA-76:000001 (US) EDB-76:000015

800	Subject Descriptor Count	Registers a count of the descriptors. In a sense, field 800 is a control field and, if present, will contain a 4-digit decimal number representing a count of descriptors, recorded in data elements 801 and 802.
801	Subject Descriptors, General	Subject descriptors which logically relate will be assigned in "groups" or "splits," and each split will be assigned a separate tag. General descriptors apply to the entire entry. The splits will be assigned tag 802 and descriptors within a specific split logically coordinate only within the split or general group, 801. Tag 801 may not be present in which case there are no general descriptors. Descriptors will be alphabetically ordered within a split (or tag) and separated by the Unit Separator (US) code. The last descriptor within a split will not have a (US) code following it. Tag 802 will be used repetitively to designate splits.
802	Subject Descriptors, Splits	
803	Descriptor Upposted Terms	These are the unique broader descriptors for the previously assigned subject descriptors in 801 and/or 802.
810	INIS Proposed Descriptors	These are terms which are not currently in the INIS Thesaurus but are being proposed by TIC to be added to the Thesaurus. These descriptors will be found in either delimiter 801 or 802. Data element 800 Subject Descriptor Count will not register the number of proposed descriptors.
950	Abstracts	Contains a short description of the contents of the entry, i.e., the abstract. The abstract can contain up to 2,000 characters. The field is free form with upper and lower case characters.

*When an item is noted as analytic it means that the document, usually a progress report, book, etc., has been resolved into chapters, sections, or other logical units for separate abstracting and indexing.