ANSI/NISO Z39.2-1994
(Revision of ANSI Z39.2-1985)

ISSN: 1041-5653

Information Interchange Format

Abstract: This standard specifies the requirements for a generalized interchange format that will accommodate data describing all forms of material. It describes a generalized structure designed specifically for exchange of data between processing systems and not necessarily for use as a processing format within systems. It may be used for the communication of records in any media.

An American National Standard
Developed by the
National Information Standards Organization
Approved April 13, 1994 by the
American National Standards Institute

Bethesda, Maryland, U.S.A.

Published by
NISO Press
P.O. Box 1056
Bethesda, MD 20827

ISSN: 1041-5653
ISBN: 1-880124-08-4

Printed in the United States of America

This paper meets the requirements of ANSI/NISO Z39.48-1992 (Permanence of Paper).

Library of Congress Cataloging-in-Publication Data

National Information Standards Organization (U.S.)
Information interchange format / developed by the National Information Standards Organization.
p. cm. — (National information standards series, ISSN 1041-5653)
"Approved . . . by the American National Standards Institute."
"ANSI/NISO Z39.2-1994."
ISBN 1-880124-08-4
1. Machine-readable bibliographic data—Standards—United States. 2. Exchange of bibliographic information—Standards—United States. 3. Museum registration methods—Data processing—Standards—United States. I. American National Standards Institute. II. Title. III. Series.
Z699.35.M28N38 1994 94-9641
025.3'16—dc20 CIP

Contents

Foreword iv

1. Scope and Field of Application 1

2. Referenced American National Standards 1

3. Definitions 1

4. Information Interchange Format 2
   4.1. Schematic Representation 2
   4.2. Leader 2
      4.2.1. Record Length 2
      4.2.2. Status 3
      4.2.3. Type of Record 3
      4.2.4. Implementation-Defined Positions 3
      4.2.5. Indicator Count 3
      4.2.6. Identifier Length 3
      4.2.7. Base Address of Data 3
      4.2.8. Implementation-Defined Positions 3
      4.2.9. Entry Map 3
   4.3. Directory 4
      4.3.1. Entries 4
         4.3.1.1. Tag 4
         4.3.1.2. Length of Field 4
         4.3.1.3. Starting Character Position 5
         4.3.1.4. Implementation-Defined Portion 5
      4.3.2. Sequence of Entries 5
   4.4. Variable Fields 5
      4.4.1. Sequence of Variable Fields 5
      4.4.2. Control Fields 5
      4.4.3. Data Fields 5
         4.4.3.1. Indicators 6
         4.4.3.2. Data Element Identifiers 6
         4.4.3.3. Multiple Data Elements 6
   4.5. Record Terminator 6

5. Other Related American National Standards 6

Figures
   Figure 1 Schematic Representation of the Information Interchange Format 3
   Figure 2 Schematic Representation of the Leader 3
   Figure 3 Entry Map 4
   Figure 4 Structure of Each Entry in the Directory 4
   Figure 5 Structure of a Variable Data Field 5

Foreword

(This foreword is not part of the American National Standard for Information Interchange Format, ANSI/NISO Z39.2-1994. It is included for information only.)

Z39.2 is one of the standards that is critical to the implementation of automation for storing, transferring, and retrieving bibliographic information. It was first issued in 1971. This 1994 edition introduces two changes: (1) removal of restrictions on character positions 07-09 in the record leader; and (2) a reduction of references to "bibliographic" data because the standard is used for many other types of data, even within the information community.

This standard was originally prepared in 1966 by Subcommittee 2 on Machine Input Records of the American National Standards Committee Z39. The Subcommittee on Machine Input Records was organized in February 1966 to define a format for the interchange of bibliographic records on magnetic tape. In its work, the subcommittee sought to ensure that: (1) the format would be hospitable to all kinds of bibliographic information and related data; (2) the format would be hardware independent; (3) the structure of all records would be basically identical and would include sufficient control information to specify unique characteristics; and (4) methods of recording and identifying data would provide for maximum manipulability leading to ease of conversion to other formats. The completed standard, ANSI Z39.2-1971, was approved by the Z39 Standards Committee in March 1969 and by the American National Standards Institute on July 14, 1970.

The Subcommittee on Machine Input Records was reconvened in 1974 for the purpose of conducting the required periodic review and revision of this standard. The principal change introduced in that revision, approved as ANSI Z39.2-1979, was the addition of an optional user-defined portion to the directory entry and the concomitant withdrawal of the requirement that entries have a length of 12 characters. This change was motivated by consideration of the problem of expressing relationships within a record. Redefinition of the directory entry did not render nonstandard any existing implementation of ANSI Z39.2-1971.

The other changes made to the format in 1979 included: (1) the reservation of additional tags for control fields; (2) the restriction of characters that may be used in tags to alphanumeric characters; (3) the specification of characters for the delimiter, the field terminator, and the record terminator (these specifications were given in Appendix A1 of ANSI Z39.2-1971); and (4) the use of a field terminator and a record terminator after the final variable field. The appendixes on implementations that were published with the original version of the standard were deleted in 1979 as unnecessary because of the appearance of extensive documentation for various implementations since the standard was first approved.

In 1985, in response to a review of the standard, the following changes were made: (1) removal of the restriction on the use of 003; and (2) expansion of the possible leader and data element identifier code values from ASCII basic characters to any ASCII graphic character (described in American National Standard Code for Information Interchange, ANSI X3.4-1977). These changes brought the standard into conformity with the existing International Standard ISO 2709, Documentation-Format for Bibliographic Information Interchange on Magnetic Tape.

NISO acknowledges with thanks and appreciation the contributions of Sally H. McCallum, Chief, Network Development and MARC Standards Office, Library of Congress, in revising this standard.

Suggestions for improving this standard are welcome. They should be sent to the National Information Standards Organization, P.O. Box 1056, Bethesda, MD 20827, telephone (301) 975-2814.

This standard was processed and approved for submittal to ANSI by the National Information Standards Organization. NISO approval of this standard does not necessarily imply that all Voting Members voted for its approval. At the time it approved this standard, NISO had the following members:

(continued)

NISO Voting Members

American Association of Law Libraries
Gary J. Bravy

American Chemical Society
Robert S. Tannehill, Jr.
Leon R. Blauvelt (Alt)

American Library Association
Jean Hamrick
Sally H. McCallum (Alt)

American Psychological Association
Lois Granick
Maurine F. Jackson (Alt)

American Society for Information Science
Clifford Lynch

American Society of Indexers
Jessica Milstead
Nancy Mulvany

American Theological Library Association
J. Raymond Vandegrift

Apple Computer, Inc.
Mike Lawler

Art Libraries Society of North America
Patricia J. Barnett
Pamela J. Parry (Alt)

Association of American Publishers (AAP)
Sandra K. Paul
Thomas D. McKee (Alt)

Association of American University Presses
Mary Lou Menches

Association of Information and Dissemination Centers
Bruce H. Kiesel

Association for Information and Image Management
Marilyn Courtot

Association of Jewish Libraries
Bella Hass Weinberg
Pearl Berger (Alt)

The Association for Recorded Sound Collections
Gerald D. Gibson
Donald McCormick (Alt)

Association of Research Libraries
Duane Webster

AT&T Bell Labs
W. David Penniman
Audrey M. Jackson (Alt)

Book Manufacturers' Institute
Douglas Horner

CAPCON Library Network
Dennis Reynolds

Catholic Library Association
Natalie A. Logan

CLSI, Inc.
Bob Walton
Andy Lukes

Colorado Alliance of Research Libraries (CARL)
Ward Shaw

Data Research Associates, Inc.
Michael J. Mellinger
James Michael (Alt)

Dynix
Rick Wilson

EBSCONET
John Krontiras
Sharon Cline McKay (Alt)

Engineering Information Inc.
Eric Johnson
Mary Berger (Alt)

The Faxon Co., Inc..
Richard Rowe
Fritz Schwartz

Gaylord Information Systems
Robert Riley
Mary Ghikas (Alt)

Indiana Cooperative Library Services Authority
Barbara Evans Markuson
Janice Cox (Alt)

Information Workstation Group
John C. Gale

Library Binding Institute
Sally Grauer

Library of Congress
Henriette D. Avram
Sally H. McCallum (Alt)

Medical Library Association
Suzanne Grefsheim

MINITEX
Anita Anker Branin
William DeJohn (Alt)

Music Library Association
Lenore Coral
Geraldine Ostrove (Alt)

National Agricultural Library
Joseph H. Howard
Samuel T. Waters (Alt)

National Archives and Records Administration
Alan Calmes

National Federation of Abstracting and Information Services
Betty Unruh
John Schnepp (Alt)

National Institute of Standards and Technology, Research Information Center, Office of Information Services
Diane Cunningham
Marietta Nelson (Alt)

National Library of Medicine
Lois Ann Colaianni

National Technical Information Service
Joseph E. Clark
John J. Elsbree (Alt)

OCLC,Inc.
Kate Nevins
Don Muccino

OHIONET
Joel Kent
Greg Pronevitz (Alt)

Optical Publishing Association
John Nairn
R. Bowers (Alt)

PALINET
James E. Rush

Pittsburgh Regional Library Center
H. E. Broadbent III

Reference Technology, Inc.
John Nairn

The Research Libraries Group, Inc.
Wayne Davison
Kathy Bales (Alt)

Society for Technical Communication
James Mason

Special Libraries Association
Audrey N. Grosch

SUNY/OCLC Network
Glyn T. Evans
David Forsythe (Alt)

UMI
Don Willis

Unisys Corporation
Michel Ridgeway
James Thomas (Alt)

U.S. Department of Commerce, Printing and Publishing Division
William S. Lofquist

U.S. Department of Defense, Army Library Management Office
Dorothy Fisk McDysan
Diane Zehnpfennig (Alt)

U.S. Department of Defense, Defense Technical Information Center
Barbara Lesser
Gretchen Schlag (Alt)

U.S. Department of Energy, Office of Scientific and Technical Information
Mary Hall
Nancy Hardin (Alt)

U.S. ISBN Maintenance Agency
Emery Koltay

U.S. National Commission on Libraries and Information Science
Robert Dugan

UTLAS International U.S., Inc.
Bill Mooney
Syd Jones (Alt)

VTLS
Vinod Chachra
Carl Lee

H.W. Wilson Company
George I. Lewicky

NISO Board of Directors

At the time NISO approved this standard the following individuals served on its Board of Directors:

Paul Evan Peters, Chairperson
New York Public Library

Mary Ellen Jacob, Immediate Past Chairperson

John Regazzi, Vice Chair/Chair-elect
Engineering Information

Heike Kordish, Treasurer
New York Public Library

Patricia R. Harris, Executive Director
National Information Standards Organization

Directors Representing Libraries

Lois Ann Colaianni
National Library of Medicine

Susan Vita
Library of Congress

Toni Carbo Bearman
University of Pittsburgh

Directors Representing Information Services

James E. Rush
PALINET

Charles Bourne
DIALOG

Wilhelm Bartenbach
The Foundation Center

Directors Representing Publishing

Peter J. Paulson
OCLC/Forest Press

Carol A. Risher
Association of American Publishers, Inc.

Karen Runyan
Houghton Mifflin Company

Information Interchange Format

1. Scope and Field of Application

This standard specifies the requirements for a generalized information interchange format that will accommodate many types of data, especially bibliographic description of all forms of materials and related data such as authority, holdings, circulation, etc.[1] It describes a generalized structure, a framework, designed specifically for exchange of data between processing systems and not necessarily for use as a processing format within systems. This standard does not specify the content of a record and does not, in general, assign meaning to tags, indicators, or data element identifiers. Such specifications shall be provided by particular implementations of the standard. The format may be used for the interchange of records using various communication media.

2. Referenced American National Standards

This standard is intended for use in conjunction with the following American National Standards. When these standards are superseded by a revision approved by the American National Standards Institute, Inc., the revision shall apply:

ANSI X3.4-1986 (R 1992), Coded Character Sets — 7-Bit American National Standard Codes for Information Interchange (7-Bit ASCII)

ANSI X3.27-1987 Magnetic Tape Labels and File Structure for Information Interchange

3. Definitions

Base address of data—A parameter whose value specifies the character position of the character following the field terminator of the directory, where the origin is the first character of the leader. (For example, if the directory contains two 12-character entries, the leader occupies positions 0-23; the two directory entries, positions 24-35 and 36-47; the directory field terminator, position 48. Consequently the base address of data for the record is 49.)

Character—A pattern of bits of a determined length (depending on the system) treated as a meaningful unit.

Communications format—See information interchange format.

Control field—A variable field containing parametric or other data which may be required for the processing of the record.

Control number—A character string uniquely associated with a record by the organization transmitting the record.

Data element—A defined unit of information.

Data element identifier—A code consisting of one or more characters used to identify individual data elements within a variable field.

Data field—A variable field containing data generally not intended to supply parameters for the processing of the record. See also **control field.**

Delimiter—ASCII character 1/15, which is used in this standard as an initiator, a separator, or a terminator of individual data elements within a variable field. (The notation given for the delimiter is that prescribed by ANSI X3.4, in which this character is named "unit separator.")

Directory—An index to the location of the variable fields (control and data) within a record. The directory consists of fixed fields called entries.

Entry—A field within the directory giving information about a variable field. (Within the text of this standard, the word "entry" always refers to a field in the directory. Since the word has other meanings, the phrase "directory entry" may be required for clarity in general use.)

Entry map—A set of parameters specifying the structure of the entries in the directory.

Field—A defined character string that may contain one or more data elements. See also **control field, data field, fixed field,** and **variable field.**

[1] The standard accommodates the description or identification of an entity independently or in relation to other entities. It does not specify a particular method for recording such relationships. For bibliographic data, for example, many types of relationships may exist, among them vertical or hierarchical relationships between a whole and its parts (for example, the relationship of a series to its individual volumes); horizontal relationships between versions of a work in different languages, formats, or media; and chronological relationships between issues of a work (for example, the relationship of a serial to its predecessors and successors). Relationships may also be recorded between different types of units: bibliographic, holdings, items, documents, etc.

Field terminator (FT)—ASCII character 1/14, which is used to terminate each variable field within a record. (The notation given for the field terminator is that prescribed by ANSI X3.4, in which this character is named "record separator.")

File—A collection of related records.

Fixed field—A field whose length is invariant, being determined for all occurrences by provisions of this standard or by definitions pertinent to a particular implementation. (Note: an implementation of this standard may fix the length of a variable field for that implementation.)

Format—The arrangement and identification of data in a record.

FT—See **field terminator**.

Identifier length—A parameter whose value gives the length of the delimiter plus the data element identifier(s) used within the record. (The values that the identifier-length parameter may assume are given in section 4.2.6.)

Indicator—A one-character data element that is associated with a data field and that supplies additional information about the field. When indicators are present, they are the first data elements in the field.

Indicator count—A parameter whose value is the number of indicators associated with each data field.

Information interchange format—A format for the exchange, rather than the local processing, of records.

Interchange format—See information interchange format.

Leader—A fixed field that occurs at the beginning of each record and provides parametric information for the processing of the record.

Length—A measure of the size of a data element, field, record, etc. Unless otherwise indicated, length in this standard is expressed in characters.

Parameter—A variable in a record that is given a constant value for the purpose of interpreting and processing that record.

Record—A collection of data elements, organized according to the provisions of this standard.

Record length—A parameter whose value is equal to the length of the entire record, including the record terminator.

Record terminator (RT)—ASCII character 1/13, which is used as the final character of a record, following the field terminator of the last data field. (The notation given for the record terminator is that prescribed by ANSI X3.4, in which this character is named "group separator.")

RT—See **record terminator**.

Status—A parameter indicating the relation of the record to a file (for example, new, updated, delete, etc.).

Structure—The framework of fixed and variable fields within a record.

Subfield—A data element considered as a component of a field.

Subrecord—A group of fields within a record that may be treated as a logical entity. (When a record describes more than one entity, the descriptions of individual entities may be treated as subrecords.)

Tag—A character string used to identify or label an associated variable field.

Type of record—A parameter that specifies the characteristics and defines the components of the record.

Variable field—A field whose length is determined for each occurrence by the length of data comprising that occurrence.

4. Information Interchange Format

4.1. Schematic Representation

The interchange format is schematically represented in Figure 1.

4.2. Leader

The leader shall comprise the first 24 character positions (0-23) of the record. It is shown schematically in Figure 2.

4.2.1. Record Length

The value of the record-length parameter shall occupy character positions 0-4 of the record. It shall be a five-digit decimal number equal to the length of the record, including the record length and the record terminator. In the interchange format, the maximum length of the record shall be 99,999 characters. (The record length described here is a logical record length for use by application programs and is not the means of communicating record length

Figure 1 — Schematic Representation of the Interchange Format

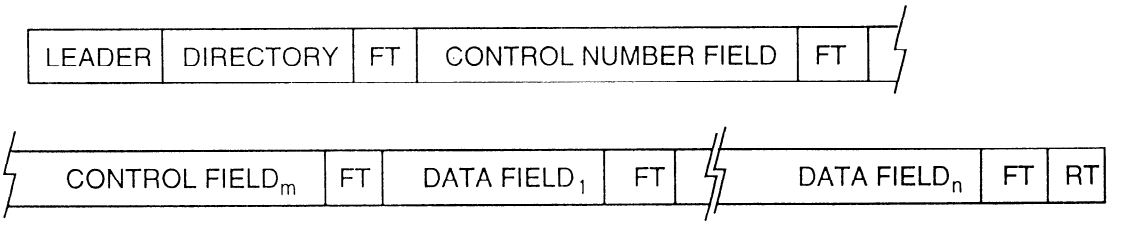
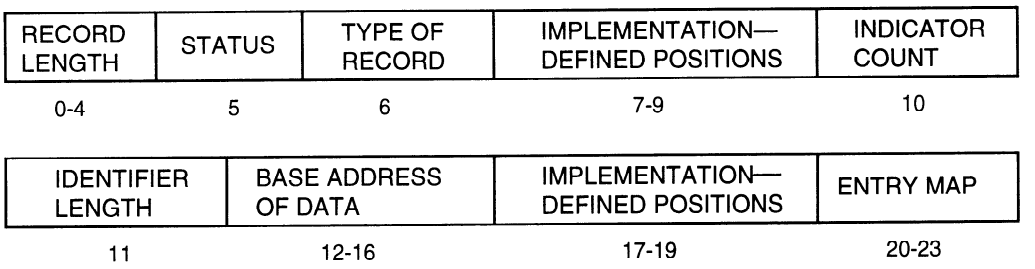


Figure 2 — Schematic Representation of the Leader



data to system software as specified in ANSI X3.27.)

4.2.2. Status

The value of the status parameter shall be an ASCII graphic character occupying character position 5 of the record.

4.2.3. Type of Record

The value of the type-of-record parameter shall be an ASCII graphic character occupying character position 6 of the record.

4.2.4. Implementation-Defined Positions

Character positions 7-9 of the record shall be reserved for definition by a particular implementation.

4.2.5. Indicator Count

The value of the indicator-count parameter shall occupy character position 10 of the record. It shall be a decimal digit equal to the number of indicators occurring in each variable data field. If the record does not include indicators, the value of this parameter shall be set to 0.

4.2.6. Identifier Length

The value of the identifier-length parameter shall occupy character position 11 of the record. It shall be a decimal digit equal to the sum of the lengths of the delimiter and the data element identifier used in the variable data fields. If the delimiter is used without a data element identifier, the value of this parameter shall be set to 1. If neither the delimiter nor a data element identifier is used, the value shall be set to 0.

4.2.7. Base Address of Data

The value of the base-address-of-data parameter shall occupy character positions 12-16 of the record. It shall be a five-digit number equal to the sum of the lengths of the leader and the directory.

4.2.8. Implementation-Defined Positions

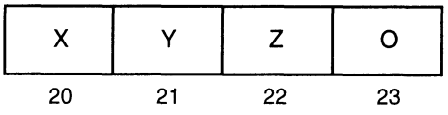
Character positions 17-19 of the record shall be reserved for definition by a particular implementation.

4.2.9. Entry Map

The entry map shall occupy character positions 20-23 of the record. (See section 4.3.1 for the description of entries.) It shall comprise four decimal

digits corresponding to the lengths of all portions of the directory entry except the tag. The first digit shall be equal to the length of the length-of-field portion of each entry. The second digit shall be equal to the length of the starting-character-position portion of the entry. The third digit shall be equal to the length of the implementation-defined portion of the entry. The fourth digit, reserved for future use, shall be set to 0. If any of these portions does not occur in the entry, the corresponding character in the entry map shall be set to 0. The entry map is shown schematically in Figure 3.

Figure 3 — Entry Map



X = length of the length-of-field portion of each directory entry
Y = length of the starting-character-position portion of each entry
Z = length of the implementation-defined portion of each entry
O = reserved for future use

The use of the entry map can be illustrated as follows: (1) an entry map set to 4500 would define the characteristics of a directory in which each entry consisted of a three-character tag (not expressed in the entry map), a four-digit length of field, and a five-digit starting character position; (2) an entry map set of 0520 would define the characteristics of a directory in which each entry consisted of a three-character tag, no length of field data element, a five-digit starting character position, and two characters of implementation-defined data.

4.3. Directory

The directory shall consist of a series of fixed fields, hereafter referred to as "entries." The directory shall contain at least one entry for each subsequent variable field (control and data). The directory shall begin in character position 24 of the record and shall end with a field terminator.

4.3.1. Entries

An entry is shown schematically in Figure 4. Each entry in the directory shall have a length equal to the sum of the lengths given in the entry map plus three characters for the tag. The tag shall be the first element of every entry. Other portions of the entry shall occur in the same sequence as do their lengths in the entry map. No portion of the entry shall exceed nine characters in length. All entries in a directory shall have the same structure. Each entry shall contain, at the very least, a tag and length of field or a tag and starting character position and shall correspond, unambiguously, to a specific variable data or control field.

4.3.1.1. Tag

A tag shall contain three alphabetic or numeric ASCII graphic characters. For alphabetic characters, an application should use either capital or lowercase letters, but not both. Tags whose first two characters are zero shall specify control fields. Other tags shall specify data fields.

For tags for control fields, tag 001 shall specify the control number field. Tag 002 shall be reserved for subrecord purposes. Tags 003-009 and 00a-00z shall be reserved for definition by particular implementations.

4.3.1.2. Length of Field

The length-of-field portion of an entry, if present, shall contain the length, expressed as a decimal number, of the variable field referenced by the entry. This length shall include the field terminator and indicators associated with the field.

Figure 4 — Structure of Each Entry in the Directory

TAG	LENGTH OF FIELD (NOTE 1)	STARTING CHARACTER POSITION (NOTE 1)	IMPLEMENTATION DEFINED PORTION (NOTE 2)

NOTES:
(1) Not required, but at least one of these portions must be present.
(2) Not required.

If the length of a variable field exceeds the maximum length expressible as a decimal number in the length-of-field portion of the entry, two or more entries (called a "subset" of the directory entries for the purposes of this explanation) shall be used to define the location and extent of such a field. All entries in the subset shall contain the same tag. The length of field in each entry of the subset, except the last entry in the subset, shall be set to 0 to indicate (a) that the length of field is equal to the maximum length expressible and (b) that there is additional information for the same field in the next entry in the record directory. The length of field for the final entry in the subset shall refer to the length of the remainder from the previous entries.

4.3.1.3. Starting Character Position

The starting-character-position portion of an entry, if present, shall contain the character position relative to the base address of data of the first character in the variable field referenced by the entry. (The first character of the first field following the directory is numbered 0). The starting character position shall be expressed as a decimal number.

4.3.1.4. Implementation-Defined Portion

The implementation-defined portion of the entry, if present, shall contain information relative to the variable field referenced by the entry. This portion of the entry shall be the same length and definition for each entry in the directory. This is not a data area.

4.3.2. Sequence of Entries

Directory entries for control fields shall precede entries for data fields. Entries for control fields shall be sequenced by tags as follows: 001, 002, ... 009, 00a, 00b, ... 00z. Entries for data fields may be recorded in any sequence and need not be in the same sequence as the corresponding variable fields.

4.4. Variable Fields

Following the leader and the directory, the record shall be composed of variable fields.

4.4.1. Sequence of Variable Fields

Control fields shall precede data fields in the record and shall be arranged in the same sequence as the corresponding entries in the directory. Data fields may be in any sequence and need not reflect the order of corresponding entries.

4.4.2. Control Fields

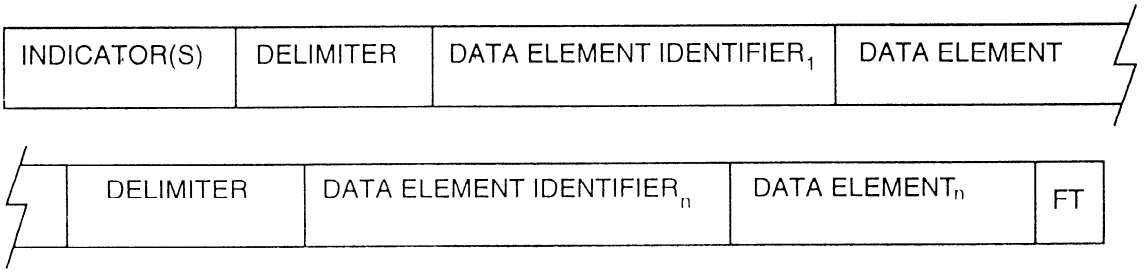
Control fields shall be specified by tags beginning with two zeroes. Control fields shall be composed of data and a field terminator; they shall not contain indicators, delimiters, or data element identifiers.

The control number field, identified by tag 001, shall contain the control number of the record. The control number shall be composed of ASCII graphic characters. Each record shall contain exactly one control number field. (As a result of the provisions of 4.4.1, the control number field will always be located at the base address of data.)

4.4.3. Data Fields

There are no restrictions on the number, length, or content of the data fields other than those already stated or implied (for example, those resulting from the limitation of total record length). Data fields may include indicators, delimiters, and data element identifiers, as well as data and a field terminator. The structure of a data field is shown in Figure 5.

Figure 5 — Structure of a Variable Data Field



4.4.3.1. Indicators

When indicators are present, they shall occupy the first character positions of the data field, preceding such delimiters and data element identifiers as may be present. Each data field in the record shall include the number of indicator character positions specified by the value of the indicator-count parameter in the leader, even if indicator values have not been defined for all fields.

4.4.3.2. Data Element Identifiers

When data element identifiers are used, each such identifier shall be preceded immediately by a delimiter and each such identifier shall immediately precede the data element it identifies. In effect, a delimiter and data element identifier are combined to form a marker used to initiate and identify data elements within a variable data field. The length of each identifier plus its associated delimiter shall be equal to the value of the identifier-length parameter in the leader and shall be uniform for each data field of a given record. If data element identifiers are used, they must precede even the first (and perhaps only) data element in the variable data fields. Data element identifiers shall be composed of ASCII graphic characters.

4.4.3.3. Multiple Data Elements

A data field may contain more than one data element. Each data element may be of fixed or variable length and may be identified by position or by the use of data element identifiers.

4.5. Record Terminator

The last character of a record shall be a record terminator following the field terminator of the last variable field.

5. Other Related American National Standards

ANSI X3.22-1983 (R 1990), Recorded Magnetic Tape for Information Interchange (800 CPI, NRZI)

ANSI X3.39-1986 (R 1992), Information Systems — Recorded Magnetic Tape for Information Interchange (1600 CPI, Phase Encoded)

ANSI X3.54-1986 (R 1992), Information Systems — Recorded Magnetic Tape for Information Interchange (6250 CPI, Group Coded Recording)

ANSI/NISO Z39.47-1993, Extended Latin Alphabet Coded Character Set for Bibliographic Use (ANSEL)

ANSI/NISO Z39.50-1992, Information Retrieval Application Service Definition and Protocol Specification for Open Systems Interconnection