

JIS

JAPANESE INDUSTRIAL STANDARD

Specification for a Data Descriptive File for Information Interchange

JIS X 0604—1986

Translated and Published

by

Japanese Standards Association

Printed in Japan

In the event of any doubt arising,
the original Standard in Japanese is to be final authority.



Documents

Document Sharing Hub

JAPANESE INDUSTRIAL STANDARD

J I S

Specification for a Data Descriptive File
for Information Interchange

X 0604-1986

1. Scope

This Japanese Industrial Standard specifies a media-independent generalized structure and its description method in regard to records and files for information interchange of data of wide varieties and structures between computer systems. Outlines of the specification contents are as follows:

- (1) A data descriptive file comprising a data descriptive record and companion data records that enable information interchange to occur with minimal external description.
- (2) The data descriptive record that describes the characteristics of each data field with the companion data records.
- (3) The description method of data elements, vectors, arrays and hierarchies containing strings, bit strings and numeric forms.
- (4) Three levels of interchange depending on the complexity of the allowed structure.

- Remarks
1. This Standard specifies the methods to describe the contents of data record, but not the contents themselves.
 2. The description method of format for numeral type data is specified in this Standard, but for the expression type of data JIS X 0210 shall be applied.
 3. The label and file structures depending on media shall be predicated on application of other specifications, such as JIS X 0601, JIS X 0602 and JIS X 0603.

2. Conformance

When the data descriptive records and all of the data records in the file for information interchange conform to the specifications of this Standard, the file shall be deemed to conform to this Standard.

A statement of conformance shall specify the interchange level to which the contents of files conform.

This Standard does not include requirements for processing of files, and therefore does not specify the conformity of the processing system of files.

Applicable Standards, Corresponding International Standard and Reference Standards: See page 41.

3. Definitions

For the purposes of this Standard the following definitions apply.

- (1) preorder traversal sequence A sequence of the nodes of a hierarchy produced by the following recursive algorithm.
 - (a) Enters the tree at the root node.
 - (b) Traverses the left-most subtree not previously traversed.
 - (c) If (b) is not possible, returns to the node superior to the subtree and goes to (b).

- (2) location The byte count indicating the position of the first byte of a field. Locations in the leader and directory are relative to the first byte of the leader, and the location of data descriptive fields and user data fields are relative to the base address of data.

- (3) alphanumeric character A character occurring in columns 2 to 7, inclusive [except position (7/15)] of the 7-bit code table for Roman letters in JIS X 0201.

Remark: The characters specified in this Standard are represented by their positions (column/row) in the coded character set table of 7 bits Roman letters in JIS X 0201, or by their acronyms, i.e., ESC, RS, US and SP. When 8-bit code system is used, the expression "(n/m) in the coded set table of 7 bits for Roman letters" is replaced by "(n/m) in the coded set table of 8 bits for Roman and Katakana letters" in JIS X 0201.

- (4) entry map A field in the leader that is used to indicate the structure of the entries in the directory.
- (5) hierarchy, hierarchical structure A rooted ordered tree structure comprising a superior root node with successive multiple ordered subtrees at increasingly inferior nodes, ultimately terminating in leaf nodes.
- (6) escape character (ESC) A control character which is used to provide additional characters. It alters the meaning of a limited number of contiguously following bit combinations. The uses shall be based on JIS X 0202.
- (7) variable-length field A field the length of which varies from occurrence to occurrence.
- (8) elementary Having the property not to be divided without loss of meaning.
- (9) null Pertaining to the condition of nonexistence of an entity, usually a data element, string or set.
- (10) delimited structure A structure composed of a collection of data elements that are separated by delimiters.

- (11) delimiter A single character that separates data elements and data fields.
- (12) unit terminator (UT) A character used to delimit several types of subfields within variable-length fields in both DDR and DR.
The character of (1/15) in the coded set table of 7 bits for Roman letters in JIS X 0201.
- (13) to map To establish the correspondence between the elements of two structures.
- (14) interchange format A format for the exchange, as opposed to the local processing, of records.
- (15) interchange level The designation of a prescribed subset of the requirements of this Standard.
- (16) relative position (RP) The position of a byte expressed as a decimal integer relative to the beginning of a field. The first relative position is numbered "0".
- (17) tag An identifier in a directory entry used to specify the internal name of an associated field.
- (18) cartesian label An array of identifiers formed by the cartesian product of the elements of two (or more) vector labels. The array elements have the same order as the elements of the direct product, so that if $\vec{a}$ and $\vec{b}$ are the vector labels, $\vec{a} = [a(1), \dots, a(n)]$ and $\vec{b} = [b(1), \dots, b(m)]$, then the cartesian label, $\vec{a} \cdot \vec{b} = [a(1) b(1), a(1) b(2), \dots, a(1) b(m), \dots, a(n) b(m)]$ where $a(i) b(j)$ is a concatenation of $a(i)$ and $b(j)$ which forms an identifier of the (i, j) element of a corresponding data array.
- (19) directory A table of directory entry as an element.
- (20) directory entry A fixed length field within the directory that contains information about the tag, the length and the location of a specific field for a given record.
- (21) data descriptive file (DDF) A file containing a data descriptive record and its companion data records.
- (22) data descriptive record (DDR) A record that logically precedes the data records and contains the control parameters and data definitions necessary to interpret companion data records. The data descriptive record is the first logical record of a file excluding the file labels.
- (23) base address of data A data element the value of which is equal to the byte count up to the first data field following the field terminator of the directory.
- (24) data record (DR) A logical record containing user data.
- (25) byte A collection of n bits.

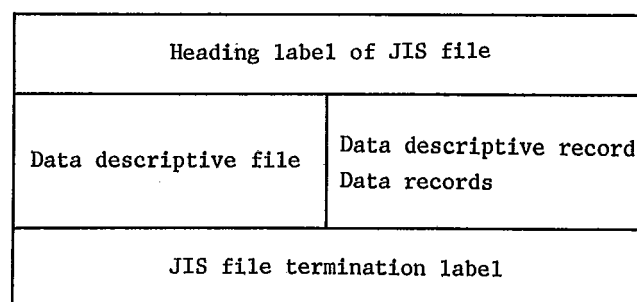
Remark: This Standard is media-independent and the number of bits is media-dependent.

- (26) array descriptor A sequence of numbers which specifies the number of dimensions and extent of an array.
- (27) bit field A data field comprising only binary digits and, when necessary, filled on the right with binary zeros to complete a byte [see also (33) character mode bit string].
- (28) file A collection of related records treated as a unit.
- (29) file title A string of characters that provides a displayable descriptive title for the interchange file. This is not necessarily the same as the file-identifying name in the label prescribed in the relevant specification for the structure of the label and file.
- (30) field terminator (FT) A character used to terminate a field within a record. The character of (1/14) in the coded set table of 7 bits for Roman letters in JIS X 0201.
- (31) compound data field A field comprising one or more elementary data elements.
- (32) vector label A vector the elements of which are labels (headings of column or row) used to identify each element in a vector of data elements.
- (33) character mode bit string A sequence of characters "0" and "1" that represents a string of binary digits [see (27) bit field].
- (34) label A character string used to identify or name the contents of subfield.
- (35) leader A fixed length field that occurs at the beginning of each record for giving parameters for the processing of the record.
- (36) record length A data element the value of which is equal to the length in bytes of the record.
- (37) logical record, record A collection of related data elements independent of their representation on a medium.

4. Interchange File

4.1 General Structures This subclause specifies the general structures of the interchange file and 4.2 to 4.3 provide the more detailed specifications. Fig. 1 shows a schematic representation of a file and the file labels.

Fig. 1. Structure of File and File Labels



This Standard specifies multiple Data Descriptive Files (DDF), each comprising logical records and the required JIS interchange file labels for the particular medium. Each DDF consists of the following logical records:

- (1) The Data Descriptive Record (DDR)
- (2) The Data Records (DR)

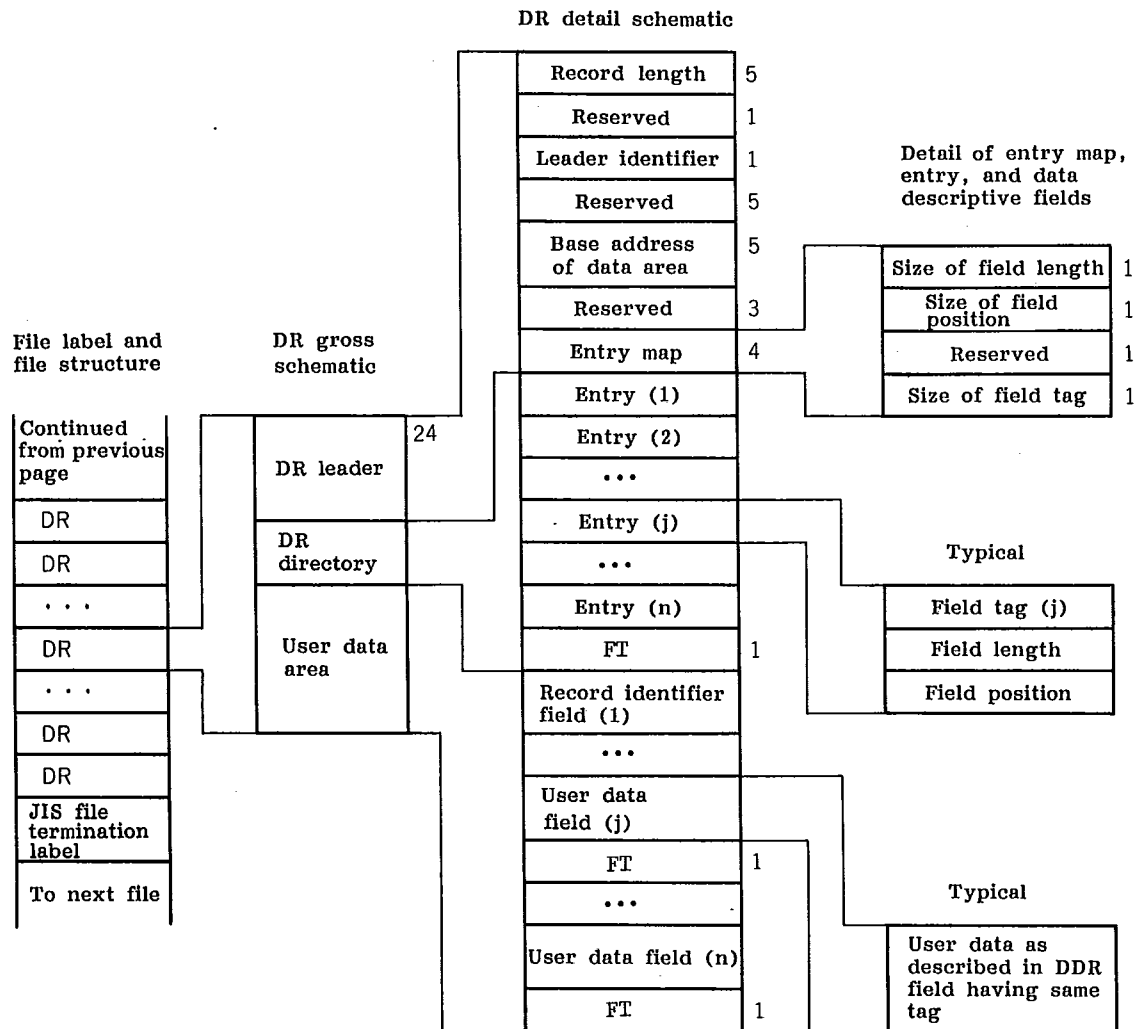
The overall structures are shown in Fig. 2, which is an expanded logical schematic representation of the DDR and DR, the leaders and directories of each record typical records and each typical data field. The DDR and DR records have the same leader, directory, field and record structure although their contents differ. A provision may be made to omit the repetition of identical DR leaders and directories for the interchange for repetitive fixed format data.

- Remarks 1. Fig. 2 shows the juxta-positions of fields, the meaning of some pointers and field lengths, and further the physical order for a medium having physically sequential access characteristic.
2. The special field tags specified in this Standard are described in the format "0 ... n", where "n" is a decimal nubmer, and "0 ..." implies sufficient zeros on the left to fill the tag field.

File label and file structure				DDR gross schematic				DDR detail schematic			
Heading label of JIS file	DDR	DR (Continued to next page)	24	DDR leader	24	Record length	5	Detail of entry map, entry, and data descriptive fields	Size of field length	1	
						Interchange level	1				
						Leader identifier	1				
						Inline code extension indicator	1				
						Reserved	1				
						Application indicator	1				
						Field control length	2				
						Base address of data descriptive area	5				
						Extended character set indicator	3				
						Entry map	4				
Entry (0)	Entry (1)	...	Entry (j)	...	Entry (n)	FT	1	Examples for level 1 file	File title		
										File control field (0)	
										FT	1
										Data description of record identifier field (1)	
										FT	1
										...	
										Data descriptive field (j)	
										FT	1
										...	
										Data descriptive field (n)	
FT	1										

- AVA** Documents
Document Sharing Hub

Fig. 2 (Cont'd)



Each data descriptive field of the DDR contains a data description of the user data in a DR user data field having the same tag. The DDR (but not the DR) has a special tag "0...0" and a corresponding field which contains field controls, an optional file title and, in the case of a hierarchy, structure information. The DRs have a special field to identify a record and the DDR contains a description of this field in the data descriptive field having the same tag "0...1". The contents of a DDR variable length field vary depending upon the values of parameters in the DDR leader.

The DDR data descriptive fields shown in Fig. 2 are for elementary character data fields and those shown in Fig. 4 are for compound data fields with all optional subfields included.

Remarks 1. The contents of a DR user data field may vary depending upon its description in the DDR, but no example is given in Figs. 2 and 4. See Reference 2 for examples of data fields.

2. Hereafter, the length of fields, except for bit fields, is given in bytes whose length in bits may be media-dependent. The contents of fields are referred to as characters, and multiple byte character sets are permitted in user fields (see 6.), and in those cases the field length is not equal to the number of characters.

8
X 0604-1986

4.2 Data Descriptive Records (DDR) The DDR shall consist of the areas and terminators shown in Table 1, and shall be the first record of the file.

Table 1. DDR Schematic

Name of area	Length
DDR leader	24
DDR directory	$k \times p$
Field terminator	1
Data descriptive area	Variable
Field terminator	1

Remark: k : Number of directory entry

p : Byte number of each entry

4.2.1 DDR Leader The DDR leader consists of the fields shown in Table 2, and the details are specified in 4.2.1.1 to 4.2.1.9.

Table 2. DDR Leader Schematic

RP ⁽¹⁾	Name of field	Length	Contents
0	Record length	5	Digits
5	Interchange level	1	Digit
6	Leader identifier	1	Character
7	Inline code extension indicator	1	Character
8	Reserved for future standardization	1	SPACE character
9	Application indicator	1	Character
10	Field control length	2	Digits
12	Base address of data	5	Digits
17	Code character set indicator	3	Characters
20	Entry map	4	Digits

Note ⁽¹⁾ RP represents the relative position [refer to 3. (16)].



4.2.1.1 Record Length Field (DDR RP 0 to 4) The total length of DDR shall be the field specified in bytes and digits. A DDR length of "0" signifies a length in excess of "99999".

4.2.1.2 Interchange Level Field (DDR RP 5) This field specifies the level of the interchange file using the following digits.

Digit "1": when the file conforms to a level 1 file.
Digit "2": when the file conforms to a level 2 file.
Digit "3": when the file conforms to a level 3 file.

A level 1 file contains elementary character data fields (see 5.1), but not compound data fields nor hierarchical structures.

A level 2 file contains compound data fields (see 5.2), but not hierarchical structures.

A level 3 file contains compound data fields and a list of the tag pairs (see 4.2.3.1.3) describing the hierarchical structures.

4.2.1.3 DDR Leader Identifier Field (DDR RP 6) This field shall designate this record as the DDR and shall contain the character "L".

4.2.1.4 Inline Code Extension Indicator (DDR RP 7) The field shall indicate whether or not inline escape sequences are used in data fields to designate extended coded character sets as specified in JIS X 0201 by means of the following characters:

Characters SP: When inline escape sequence is not used.
Character "E": When inline escape sequence is used.

4.2.1.5 Reserved for Future Standardization (DDR RP 8) This field is reserved for future standardization, and shall contain the characters SPACE.

4.2.1.6 Application Indicator Field (DDR RP 9) This field is reserved for future standardization and shall contain the characters SPACE.

4.2.1.7 Field Control Length Field (DDR RP 10 and 11) This field shall specify the number of bytes of the data descriptive field devoted to data element type and structure codes, delimiters, and other position reserved for future standardization. The contents of this field shall be the digits "00", "03", "06" or "09" (see 4.2.3.1.1, 4.3.3.2 and 6.2 to 6.4).

4.2.1.8 Base Address Field of Data Descriptive Area (DDR RP 12 to 16) The field shall specify the position of the first data descriptive field of a DDR.

Remark: The first data descriptive field is the file control field or the record identifier field.

The contents of this field shall be digits and equal to the combined length in bytes of the leader and directory including the field terminator at the end of the directory.

4.2.1.9 Default Coded Character Set Indicator Field (DDR RP 17 to 19)

The field shall specify the use of default coded character set extensions in the file by the use of the following values:

- (1) (2/0) (2/0) (2/0) When only the character set specified by the 7-bit coded set table for Roman letters in JIS X 0201 is designated as the default for the file.
- (2) (2/0) (2/1) (2/0) When extended character sets are designated as the default for one or more data fields (see 6.2).
- (3) Truncated escape sequence When an extended character set is designated as the default for the entire file (see 6.3).

4.2.1.10 Entry Map Field (DDR RP 20 to 23) The field specifies the lengths of directory entry subfields and shall consist of the subfields shown in Table 3 (see 4.2.2).

The contents of each subfield of this field shall be designated with a single digit.

Table 3. DDR Entry Map Schematic

RP	Name of subfield	Length	Contents
20	Size of field length subfield	1	Digit
21	Size of field position subfield	1	Digit
22	Reserved for future standardization	1	Digit
23	Size of field tag subfield	1	Digit

4.2.1.10.1 Size of Field Length Subfield (DDR RP 20) This field shall specify the size in bytes of the field length subfield of the directory entries using a digit in the range of "1" to "9", inclusive.

4.2.1.10.2 Size of Field Position Subfield (DDR RP 21) This field shall specify the size in bytes of the field position subfield of the directory entries using a digit in the range of "1" to "9", inclusive.

4.2.1.10.3 Reserved for Future Standardization (DDR RP 22)
The subfield is reserved for future standardization as an extended entry map and shall be the digit "0".

4.2.1.10.4 Size of Field Tag Subfield (DDR RP 23) This field shall specify the size in bytes of the field tag subfield of the directory entries using a digit in the range of "1" to "7", inclusive.

4.2.2 DDR Directory The DDR directory consists of repeated DDR directory entries. The length of each subfield of directory entry shall be specified using the entry map in the DDR leader. The DDR directory shall contain one DDR directory entry for each data descriptive field and shall end with a field terminator (1/14).

All tags appearing in the DR shall be defined in the DDR.

The DDR directory entry specifies the location and length of a corresponding data descriptive field and shall consist of the subfields shown in Table 4. The DDR directory entries shall be one to one corresponding to the data descriptive fields.

In the case of hierarchical data structures (see 4.2.3.1.3), the directory entries of the DDR shall be in the same order as the pre-order traversal sequence of the generic data tree.

Table 4. DDR Directory Entry Schematic

RP	Name of field	Length	Contents
$p \times (i-1)$	Field tag	t	Alphanumeric
$p \times (i-1) + t$	Field length	m	Digits
$p \times (i-1) + t + m$	Field position	n	Digits

Remark: $p = t + m + n$
 $i = \text{index of the directory entry}$

4.2.2.1 Field Tag This subfield contains a field tag identifying a data descriptive field and shall consist of alphanumeric ranging 1 to 7 characters. The same field tag shall occur only once in the DDR.

- (1) Field tag "0...0" identifies the optional file control field and if present shall occur only in the DDR. The file control field is optional and may or may not occur.
- (2) Field tag "0...1" occurs once in each record and identifies the record identifier field.
- (3) If used, field tag "0...2" shall identify an optional user field of the DDR which has no corresponding data field in the DR. An implementation shall pass this field to the user for processing.

Remark: The contents of this field should be determined by the user and may be used to convey any augmented file description (for example, file attributes relevant to the interchange) or ancillary file processing controls or application information.

- (4) Field tag "0...3" to "0...9" shall be reserved for future standardization.
- (5) When present, tags "0...0" to "0...9", inclusive, shall occur first in the DDR directory and in ascending numeric order.

4.2.2.2 Field Length This subfield specifies the length in bytes of the corresponding data descriptive field. This subfield value shall contain a right-adjusted integer filled with leading "0". The length of the field shall include the field terminator.

4.2.2.3 Field Position This subfield specifies the relative position of the first byte in the field referenced by the entry. This subfield value shall contain a right-justified integer filled with leading "0", and shall be located in a position relative to the base address of the data descriptive area specified in DDR RP 12 to 16. The first byte of the first field following the directory is numbered 0.

4.2.3 Data Descriptive Area The data descriptive area contains in its data fields that information which defines and describes the corresponding data fields having the same tag of the DDR as well as control parameters for automated processing. These data descriptive field shall have a specified format determined by the DDR leader.

All fields shall end with a field terminator (1/14).

4.2.3.1 File Control Field (Tag "0...0") This data descriptive field consists of the following subfields (see Fig. 3):

- (1) Field controls (if necessary)
- (2) File title (option)
- (3) List of ordered tag pairs (described only for hierarchical records)

This field shall end with a field terminator (1/14).

Fig. 3. File Control Field Schematic

Field controls	File title	U T	List of tag pairs	F T
----------------	------------	--------	----------------------	--------

4.2.3.1.1 Field Control This subfield is present only in level 2 and level 3 interchange files and its length is specified by the field control length (DDR RP 10 and 11) (see 4.2.1.7). This control information shall not be used in the file control field and shall have the value "0" or space character (see 5.2.2).

4.2.3.1.2 File Title This subfield shall be a descriptive subfield (option) of file title following the field control subfield. The file title shall be specified with a character string as representative of an external descriptive name of the interchange file.

4.2.3.1.3 List of Tag Pairs the list of tag pairs shall describe the hierarchical structure. This subfield shall follow the file title subfield and be preceded by the subfield terminators (1/15). the tag pairs shall be determined by the structural order from root node to leaf node in the generic data structure. These tag pairs may be placed in the list in any sequence but shall be contiguous without using breaking characters. The root tag shall be the tag of the record identifier "0...1". Tags "0...2" to "0...9", inclusive, do not participate in the structure specification.