

INTERNATIONAL
STANDARD

ISO/IEC
8211

Second edition
1994-10-01

**Information technology — Specification
for a data descriptive file for information
interchange**

*Technologies de l'information — Spécifications pour fichier de données
descriptif pour l'échange d'information*



Reference number
ISO/IEC 8211:1994(E)

Contents

Foreword.....	vii
Introduction.....	viii
1 Scope.....	1
2 Normative references.....	1
3 Conformance.....	2
4 Definitions and abbreviations.....	2
4.1 Definitions.....	2
4.2 Abbreviations.....	5
5 Interchange file and logical record structure.....	6
5.1 File and logical record structure.....	7
5.1.1 Interchange logical records.....	7
5.1.2 Padding of records and media blocks.....	7
5.2 Logical records - leaders and directories.....	7
5.2.1 Logical record leader.....	8
5.2.1.1 record length field (LR RP 0-4).....	8
5.2.1.2 leader identifier field (LR RP 6).....	8
5.2.1.3 ISO/IEC 8211 version number (LR RP 8).....	8
5.2.1.4 Base address of field area (LR RP 12-16).....	8
5.2.1.5 Entry map field (LR RP 20-23).....	9
5.2.1.5.1 Size of field length field (LR RP 20).....	9
5.2.1.5.2 Size of field position field (LR RP 21).....	9
5.2.1.5.3 Reserved for future standardization (LR RP 22).....	9
5.2.1.5.4 Size of field tag field (LR RP 23).....	9
5.2.1.6 Alternate forms of counts and field positions.....	9
5.2.2 Logical record directory.....	10
5.2.2.1 Field tag field.....	10
5.2.2.2 Field length field.....	10
5.2.2.3 Field position field.....	10
5.3 Logical record field areas.....	10
5.3.1 Field area of the DDR.....	10
5.3.2 The field areas of the DRs.....	11
5.3.2.1 User data fields.....	11
5.3.2.1.1 Elementary data fields.....	11
5.3.2.1.2 Compound data fields.....	11
6 Description of user data types and structures.....	11
6.1 DDR leader fields related to data description.....	11
6.1.1 Interchange level field (DDR RP 5).....	11
6.1.2 Inline code extension indicator (DDR RP 7).....	12
6.1.3 Application indicator field (DDR RP 9).....	12
6.1.3.1 Reference to other standards.....	12
6.1.4 Field control length field (DDR RP 10-11).....	12
6.1.5 Extended character set indicator field (DDR RP 17-19).....	12
6.2 Special field tags (tags = 0...0 to 0...9).....	12
6.2.1 File control field (tag = 0...0).....	12
6.2.1.1 Field control field.....	13
6.2.1.2 External file title field.....	13
6.2.1.3 List of Field tag pairs.....	13
6.2.2 Record identifier field (tag = 0...1).....	13
6.2.3 User application field (tag 0...2).....	13
6.2.4 Announcer sequence or feature identifier field (tag 0...3).....	13
6.2.5 Fields reserved for future standardization.....	13
6.2.6 Recursive tree LINKS field (tag = 0...9).....	14
6.2.7 Order of special field tags in the DDR.....	14
6.3 Data descriptive fields in level 1 files.....	14

6.4 Data descriptive fields in level 2 and 3 files.....	14
6.4.1 Tabular summary of data descriptive fields of level 2 and 3 files.....	14
6.4.2 Field controls.....	15
6.4.2.1 Data structure code (RP 0).....	16
6.4.2.2 Data type code (RP 1).....	16
6.4.2.3 Auxiliary controls (RP 2-3).....	16
6.4.2.4 Printable graphics (RP 4-5).....	16
6.4.2.5 Truncated escape sequence (RP 6-8).....	16
6.4.3 Data field names, array descriptors and format controls.....	16
6.4.3.1 Data field name.....	16
6.4.3.2 Array descriptors.....	16
6.4.3.2.1 Numeric array descriptor.....	17
6.4.3.2.2 Subfield labels	17
6.4.3.2.3 Vector labels.....	17
6.4.3.2.4 Cartesian label.....	17
6.4.3.2.5 Description of concatenated structures.....	17
6.4.3.3 Format controls.....	19
6.4.4 Order of array descriptors, labels and arrays.....	21
6.4.4.1 Order of numeric array descriptors	21
6.4.4.2 Order of cartesian labels	21
6.4.4.3 Storage order of array elements	21
7 Use of coded character sets	22
7.1 Announcement of coded character set extension	22
7.1.1 Scope of active character sets	22
7.1.2 Length of fields and subfields.....	22
7.1.3 Use of multiple octet character sets	23
7.2 ISO 2022 coded character set extension.....	23
7.2.1 Designation of ISO 2022 coded character sets.....	23
7.2.1.1 Use in the 7-bit environment.....	23
7.2.2 Designation of default code set for file.....	23
7.2.3 Designation of default code sets for fields	23
7.2.4 ISO 2022 announcer sequence field (tag 0...3).....	24
7.3 ISO/IEC 10646 coded character sets.....	24
7.3.1 Announcement of filewise default character set.....	24
7.3.2 Announcement of fieldwise default character set	24
7.3.3 ISO/IEC 10646 feature identifier field (tag 0...3)	24
Annex A ASN.1 and FTAM Registrations	26
A.1 Abstract syntax identifier.....	26
A.2 Transfer syntax identifier.....	26
A.3 FTAM document type definitions	26
A.3.1 ISO DDF unstructured document type	26
A.3.1.1 Entry number: DDF-1.....	26
A.3.1.2 Information objects	26
A.3.1.3 Scope and field of application.....	26
A.3.1.4 References.....	26
A.3.1.5 Definitions.....	26
A.3.1.6 Abbreviations	27
A.3.1.7 Document semantics	27
A.3.1.8 Abstract syntactic structure	27
A.3.1.9 Definition of transfer	27
A.3.1.9.1 Datatype definition	27
A.3.1.9.2 Presentation data values	27
A.3.1.9.3 Sequence of presentation data values	27
A.3.1.10 Transfer syntax	27
A.3.1.11 ASE specific specifications	27
A.3.1.11.1 ISO 8571 - FTAM	27
A.3.1.11.2 ISO/IEC 8211 implementation support	27
A.3.1.11.2.1 The EXTEND operation	27
A.3.1.11.2.2 The REPLACE operation	28
A.3.1.11.2.3 Relaxations	28
A.3.2 ISO DDF Structured document type	28
A.3.2.1 Entry number: DDF-2	28
A.3.2.2 Information objects	28
A.3.2.3 Scope and field of application	28
A.3.2.4 References	28

A.3.2.5 Definitions.....	28
A.3.2.6 Abbreviations.....	28
A.3.2.7 Document semantics.....	29
A.3.2.8 Abstract syntactic structure.....	29
A.3.2.9 Definition of transfer.....	29
A.3.2.9.1 Datatype definition.....	29
A.3.2.9.2 Presentation data values.....	29
A.3.2.9.3 Sequence of presentation data values.....	29
A.3.2.10 Transfer syntax.....	29
A.3.2.11 ASE specific specifications.....	30
A.3.2.11.1 ISO 8571 - FTAM.....	30
A.3.2.11.2 ISO/IEC 8211 implementation support.....	30
A.3.2.11.2.1 The EXTEND operation.....	30
A.3.2.11.2.2 The REPLACE operation.....	30
A.3.2.11.2.3 Relaxations.....	30
Annex B ISO/IEC 8211 Application Specifications.....	31
B.1 Specification of ISO/IEC 8211 Exchange File Sets.....	31
B.2 ISO/IEC 8211 data field description.....	33
B.2.1 General specifications.....	33
B.2.1.1 End of line.....	33
B.2.1.2 White space.....	33
B.2.1.3 Comments.....	33
B.2.1.4 Quoted strings.....	33
B.2.1.5 Notation.....	34
B.2.1.6 The INCLUDE construct.....	34
B.2.1.7 Order of Constructs.....	34
B.2.2 File identification.....	34
B.2.3 DDR leader specifications.....	35
B.2.4 Global default specification.....	35
B.2.5 Data field specifications.....	35
B.2.6 Special forms of field constructs.....	37
B.2.6.1 Null first vector label.....	37
B.2.6.2 Correspondence of format and last vector label.....	37
B.2.6.3 Special DDR tagged fields.....	37
B.2.7 Special constructs.....	38
B.3 Examples of exchange set specification.....	38
Annex C Informal Introduction to ISO/IEC 8211.....	42
C.1 ISO/IEC 8211 File, logical record and field constructs.....	42
C.1.1 Media record constructs.....	42
C.1.2 Logical record constructs.....	42
C.1.3 Logical record structure.....	43
C.1.3.1 Leader (RP 0 - 23).....	43
C.1.3.2 Directory.....	43
C.1.3.3 Field area.....	44
C.1.4 File characteristics and processing.....	44
C.1.5 Variant logical records.....	45
C.1.5.1 Long ISO/IEC 8211 records.....	45
C.1.5.2 Fixed-formats - repeating leaders and directories.....	45
C.1.6 ISO/IEC 8211 End-of-data conditions.....	45
C.1.7 Summary of the logical record and field constructs.....	46
C.2 Data description and identification.....	46
C.2.1 Components of data description.....	46
C.2.1.1 Data extent.....	46
C.2.1.2 Data position.....	46
C.2.1.3 Data structure.....	46
C.2.1.4 Data type and syntax.....	46
C.2.1.5 Intra-field tree structure.....	46
C.2.2 Data identification.....	47
C.2.2.1 Application semantics.....	47
C.3 File and record contents.....	47
C.4 Binary directories.....	47
Annex D Introduction to ISO/IEC 8211 Data Description.....	48
D.1 Data description - user data.....	48
D.2 Consistency of data description and data - validation.....	48
D.2.1 Complexity of data description.....	48

D.2.2 Level 1 data description.....	49
D.2.3 Level 2 and 3 data description.....	49
D.3 Data description constructs.....	50
D.3.1 Subfield extents.....	50
D.3.2 Data types.....	50
D.3.3 Field identification	50
D.3.4 Data structure without subfield identification	50
D.3.5 Data structure with subfield identification.....	51
D.4 Large application data structures	51
D.5 Intra-record tree structures	51
D.6 Coded character set extensions	51
Annex E Examples of Data Description.....	53
E.1 Leader and file title field	53
E.2 Examples of formats	54
E.2.1 Elementary data fields.....	54
E.2.2 Linear structures	54
E.2.3 Multi-dimensioned arrays	54
E.3 Examples of bit fields.....	55
E.4 Examples of binary forms	56
E.5 Examples of subfield labelling	56
E.5.1 Redundant elementary field label.....	56
E.5.2 Vector labels	56
E.5.3 Cartesian labels	57
E.5.4 Concatenated data structures	57
Annex F DDF Hierarchical and Network Data Structures	58
F.1 DDF hierarchical data structures	58
F.1.1 Forests	58
F.2 Conversion to corresponding binary tree	58
F.3 Network data structures	61
Annex G Database Data Transfer	62
G.1 Essential features of data base management systems.....	62
G.1.1 Relational data base management systems	62
G.1.2 Hierarchical data base management systems	63
G.1.3 Network data base management systems	63
G.2 Reduction to relational forms.....	63
Annex H Relationship to Other OSI Work	64
H.1 OSI basic reference model.....	64
H.1.1 Other presentation layer considerations.....	64
H.1.2 Remote versus local processing considerations	65
H.2 Relationship to FTAM virtual filestore model	65
H.2.1 Correspondence of ISO/IEC 8211 file constructs to FTAM	66
H.2.2 ISO/IEC 8211 access methodology	67
H.2.3 Relationship of documents to files	67
H.2.4 File naming	67
H.3 Relationship to other syntax notations	67
H.3.1 Abstract syntax notation one	67
H.3.2 Transfer Syntax Description Notation	68
H.4 Relationship to data base management models	68
H.5 Bibliography	68
H.6 Summary of data types in other projects	69

List of figures

Figure 1 - Schematic of ISO/IEC 8211 File and Logical Records	6
Figure 2 - File Schematic Representation	7
Figure 3 - Logical Record Schematic	7
Figure 4 - LR Leader Schematic	8
Figure 5 - LR Entry Map Schematic	9
Figure 6 - LR Directory Entry Schematic	10
Figure 7 - File Control Field Schematic	13
Figure 8 - Schematic of Level 2 and 3 Data Descriptive Fields	15
Figure F.1 - Examples of Ordered Rooted Trees	59
Figure F.2 - Generic Structure of a Logical Record	59
Figure F.3 - Instance of a User Data Tree based on F.2	60
Figure F.4 - Corresponding Binary Tree to the Tree of F.2.....	61

List of tables

Table 1 - Delimiters and Their Uses	15
Table 2 - Data Descriptive Field Components	15
Table 3 - Extensions of Bitfield Data Descriptions	20
Table A.1 Information Objects in the Unstructured Text Document Type	26
Table A.2 Information Objects in the Structured Text Document Type	28

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work.

In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC 1. Draft International Standards adopted by the joint technical committee are circulated to the national bodies for voting. Publication as an International Standard requires approval by at least 75 % of the national bodies casting a vote.

International Standard ISO/IEC 8211 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 21, *Open systems interconnection, data management and open distributed processing*.

This second edition cancels and replaces the first edition (ISO 8211:1985), which has been technically revised.

The substantive changes made to produce this edition of ISO/IEC 8211 are the following additions:

1. Binary forms for numeric values.
2. Binary leaders and directories.
3. Support for ISO/IEC 10646.
4. Definition of FTAM unstructured and structured document types.
5. Concatenated regular structures.
6. Recursive tree description.
7. A human-readable, alternate form of data field description.

The second edition is backwards compatible with the first edition.

Annex A forms an integral part of this International Standard. Annexes B to H are for information only.

Introduction

This International Standard has been produced in response to an identified need for a mechanism to allow data structures to be easily moved from one computer system to another, independent of architecture. Data structures required to be interchanged can vary significantly in complexity and size, and a common method to accomplish these interchanges is desirable. It is also desirable that any medium such as a communication line, a magnetic tape, a disk pack, a flexible disk etc., should be able to be used for the physical interchange, and that all information necessary to successfully recreate the structure in the target system should be contained within the information transported on the medium.

To meet these needs this International Standard specifies medium-independent and system-independent file and data record formats for the interchange of information between computer systems. This International Standard is intended for use with physical recorded media as well as with communications media. The contents in the user data structure can be of any internationally recognized character set and coding and are interchanged in a transparent fashion. The intermediate structure through which the information passes is designed for interchange purposes. It can also be used for some forms of general processing and is amenable to direct access methods on high volume, direct access interchange media.

This International Standard is a concrete transfer syntax and encoding standard and provides a tool for the description of files containing user data but does not specify the content or order of user data fields or user data records. It does specify a comprehensive generic form for such records and fields which can accommodate a wide variety of user needs for both simple and complex user data. An application must design its own instance of a conforming interchange file and all conforming files, both data and data description, will be processable to the field or subfield level by the same software. A user must, of course, complete the interface to their own application system.

The approach used is to define an interchange format into which most information structures and their content can be transformed without loss of information, and from which the original structure and content can be retrieved. The interchange format is suitable both for recording on physical media and transport through a communication system.

The data structures supported by the interchange format are elementary data, vectors, arrays and hierarchies. The file structures that can be transformed into the interchange format include sequential, hierarchical and relational. Network structures are not directly supported and additional pre-processing and post-processing are necessary in this case to preserve logical linkages.

This International Standard is media independent. It assumes, at a minimum, that the supporting transport system can process fixed length octet strings. It requires a computer processing capability to map the user file or database data to the interchange file. This mapping function has to provide the necessary data and structure conversions. The parameters required to define the selection and conversion of these data items and structures into the formats specified by this International Standard are outside the scope of the standard. This International Standard requires the use of a basic character set based on ISO 8859-1 and ISO/IEC 6429 in control fields and permits the use of additional character sets in user data fields. This International Standard provides for three interchange levels from which the users may choose based on the complexity of their data structures. The first interchange level supports multiple fields containing simple, unstructured character strings. The second level supports the first level and multiple fields containing structured user data comprising a variety of data types. The third level supports the second level and hierarchical data structures.

The experience of implementing ISO/IEC 8211 for a variety of applications revealed the need for the changes introduced in this version. Many of the changes give ISO/IEC 8211 increased versatility and more effective interchange capabilities. Many other changes were made to improve clarity and user acceptance. Technical changes in the standardized interchange supporting this International Standard and changes in the organizational responsibility for this International Standard have led to other extensions. This version provides the user with an improved interchange tool in keeping with the user's increasing needs and well integrated into the OSI environment.

The retrieval of archived files may require the use of computer systems which are different from the original archiving systems. The operational problems are identical to those involved in the transport of files between computers at different sites and this International Standard provides a facility for this application.

ISO/IEC 8211 is based on ISO 2709 having the same record structure but different data description components. ISO 2709 based systems for file transfer and random file access had been in use since at least 1970 and their use is now extensive. The nomenclature of ISO/IEC 8211 conforms to its predecessor, ISO 2709. Its hierarchical, logical data constructs are files, variable-length records, variable-length fields and subfields. In several programming languages, the equivalent of an ISO/IEC 8211 field is a record and specific applications will transport their records as ISO/IEC 8211 fields with related records aggregated into ISO/IEC 8211 records.

Organization of the Standard

The contents of this standard are organized as follows:

- 1) Clause 5 describes the specifications of the contents of leaders, directories and field areas common to all logical records and necessary to the import of logical records and complete fields.
- 2) Clause 6 describes the specification of the data description necessary to import user data at the subfield level. Subclause 6.1 specifies further subfields in the DDR leader which contain information pertinent to data description.
- 3) Clause 7 describes the use of extended character sets.
- 4) Annex A describes the FTAM registrations.
- 5) Annex B describes a methodology for specifying ISO/IEC 8211 file design and data descriptions.
- 6) Annexes C through H provide tutorial information on the methodology. The reader may wish to read Annex C prior to studying Clause 5 and Annex D prior to studying Clause 6.

Information technology — Specification for a data descriptive file for information interchange

1 Scope

This International Standard specifies an interchange format to facilitate the moving of files or parts of files containing data records between computer systems. The interchange format is not intended as a record format for the indigenous files of any specific system but may be used for this purpose. The standard defines a generalized structure which can be used to transmit, between systems, files or records containing a wide variety of data types and data structures. It specifies the means for the description of the contents of data records but does not specify their application semantics although these semantics can be included as a part of the transmission. The interchange format may also be used to transport individual records, individual data fields or individual subfields with their description.

This International Standard specifies:

- media-independent file and data record descriptions for information interchange;
- the description of data elements, vectors, arrays and hierarchies containing character strings, bit strings and numeric forms;
- a data descriptive file composed of a data descriptive record and companion data records that enable interchange to occur with minimal specific external description;
- the data descriptive record that describes the characteristics of each data field within the companion data records;
- three levels of complexity of file and record structure;
- FTAM unstructured and structured document types

2 Normative references

The following standards contain provisions which, through reference in the text, constitute provisions of this International Standard. At the time of its publication, the editions indicated were valid. All standards are subject to revision, and parties to agreements based on this International Standard are encouraged to investigate the possibility of applying the most recent editions of the standards indicated below. Members of IEC and ISO maintain registers of currently valid International Standards.

ISO/IEC 646:1991, *Information technology - ISO 7-bit coded character set for information interchange*

ISO 2022:1986, *Information processing - ISO 7-bit and 8-bit coded character sets - Code extension techniques*

ISO 6093:1985, *Information processing - Representation of numerical values in character strings for information interchange*

ISO/IEC 6429:1992, *Information technology - Control functions for coded character sets*

ISO 8571-1:1988, *Information processing systems - Open Systems Interconnection - File Transfer, Access and Management - Part 1: General Introduction*

ISO/IEC 8824:1990, *Information technology - Open Systems Interconnection - Specification of Abstract Syntax Notation One (ASN.1)*

ISO 8859-1:1987, *Information processing - 8-bit single-byte coded graphic character sets - Part 1: Latin alphabet No. 1*

ISO/IEC 9834-2:—1993, *Information technology - Open Systems Interconnection - Procedures for operation of OSI Registration Authorities: Part 2: Registration procedures for OSI document types*

ISO/IEC 10646-1:1993, *Information technology - Universal Multiple-Octet Coded Character Set (UCS) - Part 1: Architecture and Basic Multilingual Plane*

IEC 559:1989, *Binary floating point arithmetic for microprocessor systems* (also ANSI/IEEE 754:1985(R1991))

The following document is also relevant to this International Standard:

ISO International register of coded character sets to be used with escape sequences