

AS/NZS 3654:1996  
ISO/IEC 8211:1994

Australian/New Zealand Standard<sup>®</sup>

---

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

---

STANDARDS AUSTRALIA  
✓

STANDARDS NEW ZEALAND



Documents  
✓

Document Sharing Hub

**AS/NZS 3654:1996**

---

This Joint Australian/New Zealand Standard was prepared by Joint Technical Committee IT/1, Information Systems—Interconnection. It was approved on behalf of the Council of Standards Australia on 19 December 1995 and on behalf of the Council of Standards New Zealand on 3 May 1995. It was published on 5 May 1996.

The following interests are represented on Committee IT/1:

Australian Bankers Association  
 Australian Bureau of Statistics  
 Australian Computer Society  
 Australian Computer Users Association  
 Australian Information Industry Association  
 Australian Vice Chancellors Committee  
 Department of Defence, Australia  
 Department of Industry, Science and Technology, Australia  
 Government Computing Service, New Zealand  
 Information Exchange Steering Committee, Australia  
 Institute of Information and Communication Technologies, CSIRO, Australia  
 OPTUS Communications, Australia  
 Telstra Australia  
 Telecom, New Zealand

**Review of Standards.** To keep abreast of progress in industry, Joint Australian/New Zealand Standards are subject to periodic review and are kept up to date by the issue of amendments or new editions as necessary. It is important therefore that Standards users ensure that they are in possession of the latest edition, and any amendments thereto.

Full details of all Joint Standards and related publications will be found in the Standards Australia and Standards New Zealand Catalogue of Publications; this information is supplemented each month by the magazines 'The Australian Standard' and 'Standards New Zealand', which subscribing members receive, and which give details of new publications, new editions and amendments, and of withdrawn Standards.

Suggestions for improvements to Joint Standards, addressed to the head office of either Standards Australia or Standards New Zealand, are welcomed. Notification of any inaccuracy or ambiguity found in a Joint Australian/New Zealand Standard should be made without delay in order that the matter may be investigated and appropriate action taken.

AS/NZS 3654:1996

Australian/New Zealand Standard<sup>®</sup>

---

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

---

PUBLISHED JOINTLY BY:

STANDARDS AUSTRALIA  
1 The Crescent,  
Homebush NSW 2140 Australia

STANDARDS NEW ZEALAND  
Level 10, Standards House,  
155 The Terrace,  
Wellington 6001 New Zealand

## PREFACE

This Standard was prepared by the Joint Standards Australia/Standards New Zealand Committee IT/1 on Information Systems—Interconnection. It is identical with and has been reproduced from ISO/IEC 8211:1994, *Information technology—Specification for a data descriptive file for information interchange* and will supersede AS 3654—1989, *Information processing—Specification for a data descriptive file for information interchange*.

The objective of this Standard is to provide users of multiplatform environments with a medium-independent and system independent file and data record formats for the interchange of information.

The term ‘informative’ has been used in the Standard to define the application of the annex to which it applies. An ‘informative’ annex is only for information and guidance. A ‘normative’ annex is an integral part of a Standard.

Users of this Standard are advised by Standards Australia and Standards New Zealand, under arrangements made with ISO and IEC, as well as certain other Standards organizations, that the number of this Standard is not reproduced on each page; its identity is shown only on the cover and title pages.

For the purpose of this Standard, the source text should be modified as follows:

- (a) *Terminology* The words ‘this Australian/New Zealand Standard’ should replace the words ‘this International Standard’ wherever they appear.
- (b) *Decimal marker* Substitute a full point for a comma where it appears as a decimal marker.
- (c) *References* The references to international Standards should be replaced by references, where appropriate, to the following Australian or Joint Australian/New Zealand Standards:

| <i>Reference to International Standard<br/>or other publication</i> |                                                                                                                         | <i>Australian or Joint<br/>Australian/New Zealand Standard</i> |                                                                                                                             |
|---------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------|
| IEC                                                                 |                                                                                                                         | AS                                                             |                                                                                                                             |
| 559                                                                 | Binary floating point arithmetic for microprocessor systems                                                             | —                                                              |                                                                                                                             |
| ISO/IEC                                                             |                                                                                                                         |                                                                |                                                                                                                             |
| 646                                                                 | Information technology—<br>ISO 7 bit coded character set for<br>information interchange                                 | 1776                                                           | Information processing—<br>7-bit coded character set for<br>information interchange                                         |
| 6429                                                                | Information technology—<br>Control functions for coded character<br>sets                                                | 2761                                                           | Character sets and information coding—<br>Control functions for 7-bit and 8-bit<br>coded character sets                     |
| 8824                                                                | Information technology—<br>Open Systems Interconnection—<br>Specification of Abstract Syntax<br>Notation One (ASN.1)    | 3625                                                           | Information technology—<br>Open Systems Interconnection—<br>Specification of Abstract Syntax<br>Notation One (ASN.1)        |
|                                                                     |                                                                                                                         | AS/NZS                                                         |                                                                                                                             |
| 9834                                                                | Information technology—<br>Open Systems Interconnection—<br>Procedures for operation of OSI<br>Registration Authorities | 4153                                                           | Information technology—<br>Open Systems Interconnection—<br>Procedures for the operation of OSI<br>registration authorities |
| 9834-2                                                              | Part 2: Registration procedures for<br>OSI document types                                                               | 4153.2                                                         | Part 2: Registration procedures for<br>OSI document types                                                                   |

## iii

| ISO/IEC                                                                                                                   | AS/NZS                                                                                                                    |
|---------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------|
| 10646 Information technology—<br>Universal Multiple-Octet Coded<br>Character Set (UCS)                                    | 4189 Information technology—<br>Universal multiple-octet coded<br>Character Set (UCS)                                     |
| 10646-1 Part 1: Architecture and Basic<br>Multilingual Plane                                                              | 4189.1 Part 1: Architecture and basic<br>multilingual plane                                                               |
| 2022 Information processing—<br>ISO 7-bit and 8-bit coded character<br>sets—Code extension techniques                     | AS<br>1953 Information processing—<br>ISO 7-bit and 8-bit coded character sets—<br>Code extension techniques              |
| 6093 Information processing—<br>Representation of numerical values in<br>character strings for information<br>interchange | 3767 Information processing—<br>Representation of numerical values in<br>character strings for information<br>interchange |
| ISO                                                                                                                       | AS/NZS                                                                                                                    |
| 8751 Information processing systems—<br>Open Systems Interconnection—<br>File Transfer, Access and Management—            | 3646 Information processing systems—<br>Open Systems Interconnection—<br>File transfer, access and management             |
| 8751-1 Part 1: General Introduction                                                                                       | 3646.1 Part 1: General introduction                                                                                       |
| 8859 Information processing—<br>8-bit single-byte coded graphic<br>character sets                                         | —                                                                                                                         |
| 8859-1 Part 1: Latin alphabet No.1                                                                                        |                                                                                                                           |
| ISO International register of coded<br>character sets to be used with escape<br>sequences                                 |                                                                                                                           |

## © Copyright – STANDARDS AUSTRALIA/STANDARDS NEW ZEALAND

Users of Standards are reminded that copyright subsists in all Standards Australia and Standards New Zealand publications and software. Except where the Copyright Act allows and except where provided for below no publications or software produced by Standards Australia or Standards New Zealand may be reproduced, stored in a retrieval system in any form or transmitted by any means without prior permission in writing from Standards Australia or Standards New Zealand. Permission may be conditional on an appropriate royalty payment. Australian requests for permission and information on commercial software royalties should be directed to the head office of Standards Australia. New Zealand requests should be directed to Standards New Zealand.

Up to 10 percent of the technical content pages of a Standard may be copied for use exclusively in-house by purchasers of the Standard without payment of a royalty or advice to Standards Australia or Standards New Zealand.

Inclusion of copyright material in computer software programs is also permitted without royalty payment provided such programs are used exclusively in-house by the creators of the programs.

Care should be taken to ensure that material used is from the current edition of the Standard and that it is updated whenever the Standard is amended or revised. The number and date of the Standard should therefore be clearly identified.

The use of material in print form or in computer software programs to be used commercially, with or without payment, or in commercial contracts is subject to the payment of a royalty. This policy may be varied by Standards Australia or Standards New Zealand at any time.

## CONTENTS

|                                                                       | <i>Page</i> |
|-----------------------------------------------------------------------|-------------|
| 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.....                                                 | 18 |
| 6.4.4 Order of array descriptors, labels and arrays.....                     | 20 |
| 6.4.4.1 Order of numeric array descriptors.....                              | 20 |
| 6.4.4.2 Order of cartesian labels.....                                       | 20 |
| 6.4.4.3 Storage order of array elements.....                                 | 20 |
| 7 Use of coded character sets.....                                           | 21 |
| 7.1 Announcement of coded character set extension.....                       | 21 |
| 7.1.1 Scope of active character sets.....                                    | 21 |
| 7.1.2 Length of fields and subfields.....                                    | 21 |
| 7.1.3 Use of multiple octet character sets.....                              | 22 |
| 7.2 ISO 2022 coded character set extension.....                              | 22 |
| 7.2.1 Designation of ISO 2022 coded character sets.....                      | 22 |
| 7.2.1.1 Use in the 7-bit environment.....                                    | 22 |
| 7.2.2 Designation of default code set for file.....                          | 22 |
| 7.2.3 Designation of default code sets for fields.....                       | 22 |
| 7.2.4 ISO 2022 announcer sequence field (tag 0...3).....                     | 23 |
| 7.3 ISO/IEC 10646 coded character sets.....                                  | 23 |
| 7.3.1 Announcement of filewise default character set.....                    | 23 |
| 7.3.2 Announcement of fieldwise default character set.....                   | 23 |
| 7.3.3 ISO/IEC 10646 feature identifier field (tag 0...3).....                | 23 |
| Annex A ASN.1 and FTAM Registrations.....                                    | 25 |
| A.1 Abstract syntax identifier.....                                          | 25 |
| A.2 Transfer syntax identifier.....                                          | 25 |
| A.3 FTAM document type definitions.....                                      | 25 |
| A.3.1 ISO DDF unstructured document type.....                                | 25 |
| A.3.1.1 Entry number: DDF-1.....                                             | 25 |
| A.3.1.2 Information objects.....                                             | 25 |
| A.3.1.3 Scope and field of application.....                                  | 25 |
| A.3.1.4 References.....                                                      | 25 |
| A.3.1.5 Definitions.....                                                     | 25 |
| A.3.1.6 Abbreviations.....                                                   | 26 |
| A.3.1.7 Document semantics.....                                              | 26 |
| A.3.1.8 Abstract syntactic structure.....                                    | 26 |
| A.3.1.9 Definition of transfer.....                                          | 26 |
| A.3.1.9.1 Datatype definition.....                                           | 26 |
| A.3.1.9.2 Presentation data values.....                                      | 26 |
| A.3.1.9.3 Sequence of presentation data values.....                          | 26 |
| A.3.1.10 Transfer syntax.....                                                | 26 |
| A.3.1.11 ASE specific specifications.....                                    | 26 |
| A.3.1.11.1 ISO 8571 - FTAM.....                                              | 26 |
| A.3.1.11.2 ISO/IEC 8211 implementation support.....                          | 26 |
| A.3.1.11.2.1 The EXTEND operation.....                                       | 26 |
| A.3.1.11.2.2 The REPLACE operation.....                                      | 27 |
| A.3.1.11.2.3 Relaxations.....                                                | 27 |
| A.3.2 ISO DDF Structured document type.....                                  | 27 |
| A.3.2.1 Entry number: DDF-2.....                                             | 27 |
| A.3.2.2 Information objects.....                                             | 27 |
| A.3.2.3 Scope and field of application.....                                  | 27 |
| A.3.2.4 References.....                                                      | 27 |

|                                                                 | <i>Page</i> |
|-----------------------------------------------------------------|-------------|
| A.3.2.5 Definitions.....                                        | 27          |
| A.3.2.6 Abbreviations.....                                      | 27          |
| A.3.2.7 Document semantics.....                                 | 28          |
| A.3.2.8 Abstract syntactic structure.....                       | 28          |
| A.3.2.9 Definition of transfer.....                             | 28          |
| A.3.2.9.1 Datatype definition.....                              | 28          |
| A.3.2.9.2 Presentation data values.....                         | 28          |
| A.3.2.9.3 Sequence of presentation data values.....             | 28          |
| A.3.2.10 Transfer syntax.....                                   | 28          |
| A.3.2.11 ASE specific specifications.....                       | 29          |
| A.3.2.11.1 ISO 8571 - FTAM.....                                 | 29          |
| A.3.2.11.2 ISO/IEC 8211 implementation support.....             | 29          |
| A.3.2.11.2.1 The EXTEND operation.....                          | 29          |
| A.3.2.11.2.2 The REPLACE operation.....                         | 29          |
| A.3.2.11.2.3 Relaxations.....                                   | 29          |
| Annex B ISO/IEC 8211 Application Specifications.....            | 30          |
| B.1 Specification of ISO/IEC 8211 Exchange File Sets.....       | 30          |
| B.2 ISO/IEC 8211 data field description.....                    | 32          |
| B.2.1 General specifications.....                               | 32          |
| B.2.1.1 End of line.....                                        | 32          |
| B.2.1.2 White space.....                                        | 32          |
| B.2.1.3 Comments.....                                           | 32          |
| B.2.1.4 Quoted strings.....                                     | 32          |
| B.2.1.5 Notation.....                                           | 33          |
| B.2.1.6 The INCLUDE construct.....                              | 33          |
| B.2.1.7 Order of Constructs.....                                | 33          |
| B.2.2 File identification.....                                  | 33          |
| B.2.3 DDR leader specifications.....                            | 34          |
| B.2.4 Global default specification.....                         | 34          |
| B.2.5 Data field specifications.....                            | 34          |
| B.2.6 Special forms of field constructs.....                    | 36          |
| B.2.6.1 Null first vector label.....                            | 36          |
| B.2.6.2 Correspondence of format and last vector label.....     | 36          |
| B.2.6.3 Special DDR tagged fields.....                          | 36          |
| B.2.7 Special constructs.....                                   | 37          |
| B.3 Examples of exchange set specification.....                 | 37          |
| Annex C Informal Introduction to ISO/IEC 8211.....              | 41          |
| C.1 ISO/IEC 8211 File, logical record and field constructs..... | 41          |
| C.1.1 Media record constructs.....                              | 41          |
| C.1.2 Logical record constructs.....                            | 41          |
| C.1.3 Logical record structure.....                             | 42          |
| C.1.3.1 Leader (RP 0 - 23).....                                 | 42          |
| C.1.3.2 Directory.....                                          | 42          |
| C.1.3.3 Field area.....                                         | 43          |
| C.1.4 File characteristics and processing.....                  | 43          |
| C.1.5 Variant logical records.....                              | 44          |
| C.1.5.1 Long ISO/IEC 8211 records.....                          | 44          |
| C.1.5.2 Fixed-formats - repeating leaders and directories.....  | 44          |
| C.1.6 ISO/IEC 8211 End-of-data conditions.....                  | 44          |
| C.1.7 Summary of the logical record and field constructs.....   | 45          |
| C.2 Data description and identification.....                    | 45          |
| C.2.1 Components of data description.....                       | 45          |
| C.2.1.1 Data extent.....                                        | 45          |
| C.2.1.2 Data position.....                                      | 45          |
| C.2.1.3 Data structure.....                                     | 45          |
| C.2.1.4 Data type and syntax.....                               | 45          |
| C.2.1.5 Intra-field tree structure.....                         | 45          |
| C.2.2 Data identification.....                                  | 46          |
| C.2.2.1 Application semantics.....                              | 46          |
| C.3 File and record contents.....                               | 46          |
| C.4 Binary directories.....                                     | 46          |
| Annex D Introduction to ISO/IEC 8211 Data Description.....      | 47          |
| D.1 Data description - user data.....                           | 47          |
| D.2 Consistency of data description and data - validation.....  | 47          |
| D.2.1 Complexity of data description.....                       | 47          |

|                                                                   | <i>Page</i> |
|-------------------------------------------------------------------|-------------|
| D.2.2 Level 1 data description.....                               | 48          |
| D.2.3 Level 2 and 3 data description.....                         | 48          |
| D.3 Data description constructs.....                              | 49          |
| D.3.1 Subfield extents.....                                       | 49          |
| D.3.2 Data types.....                                             | 49          |
| D.3.3 Field identification.....                                   | 49          |
| D.3.4 Data structure without subfield identification.....         | 49          |
| D.3.5 Data structure with subfield identification.....            | 50          |
| D.4 Large application data structures.....                        | 50          |
| D.5 Intra-record tree structures.....                             | 50          |
| D.6 Coded character set extensions.....                           | 50          |
| Annex E Examples of Data Description.....                         | 52          |
| E.1 Leader and file title field.....                              | 52          |
| E.2 Examples of formats.....                                      | 53          |
| E.2.1 Elementary data fields.....                                 | 53          |
| E.2.2 Linear structures.....                                      | 53          |
| E.2.3 Multi-dimensioned arrays.....                               | 53          |
| E.3 Examples of bit fields.....                                   | 54          |
| E.4 Examples of binary forms.....                                 | 55          |
| E.5 Examples of subfield labelling.....                           | 55          |
| E.5.1 Redundant elementary field label.....                       | 55          |
| E.5.2 Vector labels.....                                          | 55          |
| E.5.3 Cartesian labels.....                                       | 56          |
| E.5.4 Concatenated data structures.....                           | 56          |
| Annex F DDF Hierarchical and Network Data Structures.....         | 57          |
| F.1 DDF hierarchical data structures.....                         | 57          |
| F.1.1 Forests.....                                                | 57          |
| F.2 Conversion to corresponding binary tree.....                  | 57          |
| F.3 Network data structures.....                                  | 60          |
| Annex G Database Data Transfer.....                               | 61          |
| G.1 Essential features of data base management systems.....       | 61          |
| G.1.1 Relational data base management systems.....                | 61          |
| G.1.2 Hierarchical data base management systems.....              | 62          |
| G.1.3 Network data base management systems.....                   | 62          |
| G.2 Reduction to relational forms.....                            | 62          |
| Annex H Relationship to Other OSI Work.....                       | 63          |
| H.1 OSI basic reference model.....                                | 63          |
| H.1.1 Other presentation layer considerations.....                | 63          |
| H.1.2 Remote versus local processing considerations.....          | 64          |
| H.2 Relationship to FTAM virtual filestore model.....             | 64          |
| H.2.1 Correspondence of ISO/IEC 8211 file constructs to FTAM..... | 65          |
| H.2.2 ISO/IEC 8211 access methodology.....                        | 66          |
| H.2.3 Relationship of documents to files.....                     | 66          |
| H.2.4 File naming.....                                            | 66          |
| H.3 Relationship to other syntax notations.....                   | 66          |
| H.3.1 Abstract syntax notation one.....                           | 66          |
| H.3.2 Transfer Syntax Description Notation.....                   | 67          |
| H.4 Relationship to data base management models.....              | 67          |
| H.5 Bibliography.....                                             | 67          |
| H.6 Summary of data types in other projects.....                  | 68          |

## 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.....                 | 58 |
| Figure F.2 - Generic Structure of a Logical Record.....            | 59 |
| Figure F.3 - Instance of a User Data Tree based on F.2.....        | 59 |
| Figure F.4 - Corresponding Binary Tree to the Tree of F.2.....     | 60 |



List of tables

Table 1 - Delimiters and Their Uses ..... 15

Table 2 - Data Descriptive Field Components ..... 15

Table 3 - Extensions of Bitfield Data Descriptions ..... 19

Table A.1 Information Objects in the Unstructured Text Document Type ..... 25

Table A.2 Information Objects in the Structured Text Document Type ..... 27

Originated as AS 3654—1989.  
Jointly revised and designated AS/NZS 3654:1996.

## AUSTRALIAN/NEW ZEALAND STANDARD

**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*

### 3 Conformance

Interchange files conform to this International Standard when all of the data descriptive records and data records conform to the specifications of this International Standard. A statement of conformance shall specify the version number and interchange level to which the contents of files conform.

This International Standard does not specify requirements for processing and implementation, therefore such processing and implementation cannot itself conform to this International Standard.

### 4 Definitions and abbreviations

#### 4.1 Definitions

For the purposes of this International Standard the following definitions apply:

**4.1.1 alphanumeric character:** A character occurring in columns 2 to 7 inclusive (except position (7/15)) of the Basic Character Set or the corresponding characters of ISO/IEC 10646 Table 1 - Row 00: Basic Latin. These characters correspond to those of ISO/IEC 646 IRV.

NOTE 1 - The characters specified in this International Standard are represented by their position (column/row) in the coded character set table or by their acronym or name, e. g., ESC, SPACE and DIGIT ZERO or "0". Alphanumeric characters which are explicitly specified in a control field are enclosed in double quotes, e.g., "1". The special ISO/IEC 8211 delimiters unit terminator(1/13) and field terminator(1/14) are represented by UT and FT respectively.

**4.1.2 array:** A data structure of two or more dimensions.

**4.1.3 array descriptor:** A Cartesian label or numeric array descriptor which provides a description of the dimension and extents of an array.

**4.1.4 base address of data:** A data element the value of which is equal to the octet count up to but not including the first octet of the first data field following the field terminator of the directory, where the specified origin (0) is the first octet of the leader.

**4.1.5 basic character set; BCS:** A character set comprising: a) ISO 8859-1 including SPACE as the G0 and G1 sets and b) ISO/IEC 6429 as the C0 and C1 sets.

NOTE 2 - This set is compatible with row 00 of the ISO/IEC 10646 Basic Multilingual Plane and should greatly reduce any dependence on national variant character sets. However there currently are differences in the allowed uses of the C1 set between ISO/IEC 6429 and ISO/IEC 10646.

**4.1.6 binary form:** A binary form is a sequence of one or more octets having a specified format and meaning. The octets of a binary form may be recorded in order of decreasing significance with the most significant octet first (MSOF) or the reverse order, least significant octet first (LSOF). The bits within an octet are always recorded in order of decreasing significance.

**4.1.7 binary tree :** A rooted tree in which each node has zero subtrees or, at most, two subtrees known as the left subtree and the right subtree. In the absence of either the left or right subtree, the remaining subtree retains its identity as a left or right subtree. See annex F.

**4.1.8 bit field:** A data field comprising only binary digits. When it is necessary to complete an octet, it is filled on the right with binary zeros. See also **character mode bit string**.

**4.1.9 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 such that if  $a$  and  $b$  are the vector labels, where  $a = (a(1) \dots, a(n))$  and  $b = (b(1) \dots, b(m))$ , then the Cartesian label,  $a*b = (a(1)b(1), a(1)b(2), \dots, a(1)b(m), \dots, a(n)b(m))$ , where  $a(i)b(j)$  is a concatenation of  $a(i)$  and  $b(j)$  which forms an identifier of the  $j,i$ -th element of a corresponding data array (i.e., the "a"s are the row labels and the "b"s are the column labels). The Cartesian label is the most general form of a label and contains a vector label and a single label as special cases. The expansion of a higher order Cartesian label is from the left, i.e.,  $a*b*c = (a*b)*c$ , that is,  $a*b$  is expanded first.

**4.1.10 character mode bit string:** A sequence of alphanumeric characters, i. e. "0" or "1", that represents a string of binary digits. See also **bit field**.

**4.1.11 compound data field:** A field comprising one or more subfields each containing an elementary data element.

**4.1.12 corresponding binary tree:** A binary tree which represents the structure of an ordered, rooted tree. The first (leftmost) offspring of a node in an ordered rooted tree becomes the root of the left subtree of the binary tree and the siblings of the node are connected through a series of right-hand links. The preorder traversal sequence of an ordered, rooted tree and its corresponding binary tree are identical.

**4.1.13 data descriptive field:** A tagged field, residing in the DDR, containing the ISO/IEC 8211 data description for all data fields in the file associated with the same tag.

**4.1.14 data descriptive file; DDF:** A file containing a data descriptive record and its companion data records.

**4.1.15 data descriptive record; DDR:** A logical record that contains the control parameters and data definitions necessary to interpret its companion data records. The data descriptive record is the first logical record of a file.

**4.1.16 data field:** A tagged field which contains user data and is located in a DR.

**4.1.17 data record; DR:** A logical record containing user data.

**4.1.18 delimited structure:** A data structure composed of a collection of data elements that are separated by delimiters.

**4.1.19 delimiter:** A single character that separates data elements and data subfields. See 6.4.1, Table 1 for the use of delimiters.

**4.1.20 directory:** A field containing an array of identifiers and references to corresponding items of data. It is terminated by a field terminator.

**4.1.21 directory entry:** A fixed-length field within the directory that contains information about the field tag, the length and the location of a specific field within a given record.

**4.1.22 elementary:** Having the property of being indivisible without loss of meaning.

**4.1.23 entry map:** A field in the leader that is used to indicate the structure of the entries in the directory.

**4.1.24 escape character; ESC:** A control character which is used for code extension purposes. It causes the meaning of a limited number of bit combinations following it to be changed. The use of ESC is specified in ISO 2022.

**4.1.25 external file title:** A string of characters that provides a displayable descriptive title for the interchange file. This need not be the same as the host system file name.

**4.1.26 field:** A generic term used to describe a contiguous set of octets containing one or more values. The components of a field are referred to generically as subfields. See also **subfield 2**).

**4.1.27 field tag; tag:** A character string in a directory entry used to identify a data field or an associated data descriptive field.

**4.1.28 field terminator; FT:** A character, (1/14), used to terminate a variable-length field within a record.

**4.1.29 file:** A collection of related records treated as a unit.

**4.1.30 forest:** An ordered set of disjoint, ordered, rooted trees.

**4.1.31 interchange format:** A format for the exchange, as opposed to the local processing, of data.

**4.1.32 interchange level; level:** The designation of a prescribed subset of the requirements of this International Standard.

**4.1.33 leader:** A fixed-length field that occurs at the beginning of each record and provides parameters for the processing of that record.

**4.1.34 least significant octet first; LSOF:** An ordering of octets in which the least significant octet is placed closest to the beginning of a file.

**4.1.35 location; position:** The octet count to the position of the first octet of a field. Locations in the leader and directory are relative to the first (0) octet of the leader, and the locations of fields are relative to the base address of data.

**4.1.36 map; to map:** To establish the correspondence between the elements of two different data structures.

**4.1.37 media record:** The physical record which is characteristic of a medium and which is written on the media by a single system level write statement, e.g., a block or a sector or a transmission packet. (see also **record**)

**4.1.38 most significant octet first; MSOF:** An ordering of octets in which the most significant octet is placed closest to the beginning of a file.

**4.1.39 n-tuple:** A vector, each instance of which has n ordered elements.

**4.1.40 null:** The non-occurrence of an entity, usually a data element, string or set.

**4.1.41 numeric array descriptor:** A sequence of numbers which specifies the number of dimensions and extent of each dimension of an array.

**4.1.42 ordered, rooted tree; tree; hierarchy:** A data structure comprising a set of nodes having directed links such that one node, the root node, has no incoming links, the remaining nodes have precisely one incoming link, and the order of multiple outgoing links from a node is significant. The nodes are referred to as parent nodes and offspring nodes and link direction is from parent node to offspring node.

**4.1.43 position:** see **location**.

**4.1.44 preorder traversal sequence:** A sequence of the nodes of an ordered, rooted tree which is produced by the following recursive algorithm:

- a) Enter the tree at the root node.
- b) Traverse the left-most subtree not previously traversed,  
if no such subtree exists, stop.
- c) If b) is not possible, return to the node superior to the subtree and go to b).

**4.1.45 record; logical record:** A contiguous string of octets in an interchange file, the first five being the record length (see also **media record**).

**4.1.46 record length:** A data element the value of which is equal to the length in octets of the record.

**4.1.47 recursive tree:** A tree in which a node appears in its own subtree.

**4.1.48 relative position; RP:** The position of an octet expressed as a decimal integer relative to the beginning of a record, field or subfield. The relative position of the first octet is numbered "0."

**4.1.49 signed binary integer:** A two's complement binary number.

**4.1.50 subfield:** 1) A contiguous string of octets in a field whose position, length and data type are described in the field data description; 2) generically, a component of a field.

**4.1.51 subfield label:** A character string used in the data description to identify a subfield and its contents.

**4.1.52 tagged field:** A contiguous string of octets in a record which is identified by a field tag and whose length and relative position are specified in a directory entry. A tagged field may occur in the DDR or in the DRs.

**4.1.53 tag:** see field tag.

**4.1.54 unit terminator; UT:** A character, (1/15), used to delimit several types of subfields within variable-length fields in both DDR and DR.

**4.1.55 variable-length field:** A field, the length of which may vary from occurrence to occurrence.

**4.1.56 vector:** A one dimensional data structure each instance of which may have a variable number of elements. (see also n-tuple)

**4.1.57 vector label:** A vector the elements of which are labels used to identify each element in a vector of data elements. A vector label is a special case of a Cartesian label.

## 4.2 Abbreviations

For the purposes of this International Standard the following abbreviations apply:

|       |                                       |
|-------|---------------------------------------|
| BCS   | basic character set                   |
| DDF   | data descriptive file                 |
| DDR   | data descriptive record               |
| DFD   | data field description, see B.2       |
| DR    | data record                           |
| ESC   | escape character                      |
| FT    | field terminator                      |
| FTAM  | file transfer and access management   |
| LR    | logical record                        |
| LSOF  | least significant octet first         |
| MSBF  | most significant octet first          |
| RP    | relative position                     |
| SPACE | SPACE character                       |
| TAB   | horizontal tabulation character       |
| UT    | unit terminator                       |
| level | interchange level                     |
| tag   | field tag                             |
| 0...n | a tag with zero or more leading ZEROs |

## 5 Interchange file and logical record structure

This clause specifies the general requirements for an interchange file and the requirements which are common to all logical records. These requirements specify the structure of logical records external to the fields. The requirements for the contents of the fields and the description of the contents of data fields are given in clause 6. See Annex C for an informal description of the constructs of this clause.

An introductory schematic for ISO/IEC 8211 files, logical records and their components is shown in Figure 1.

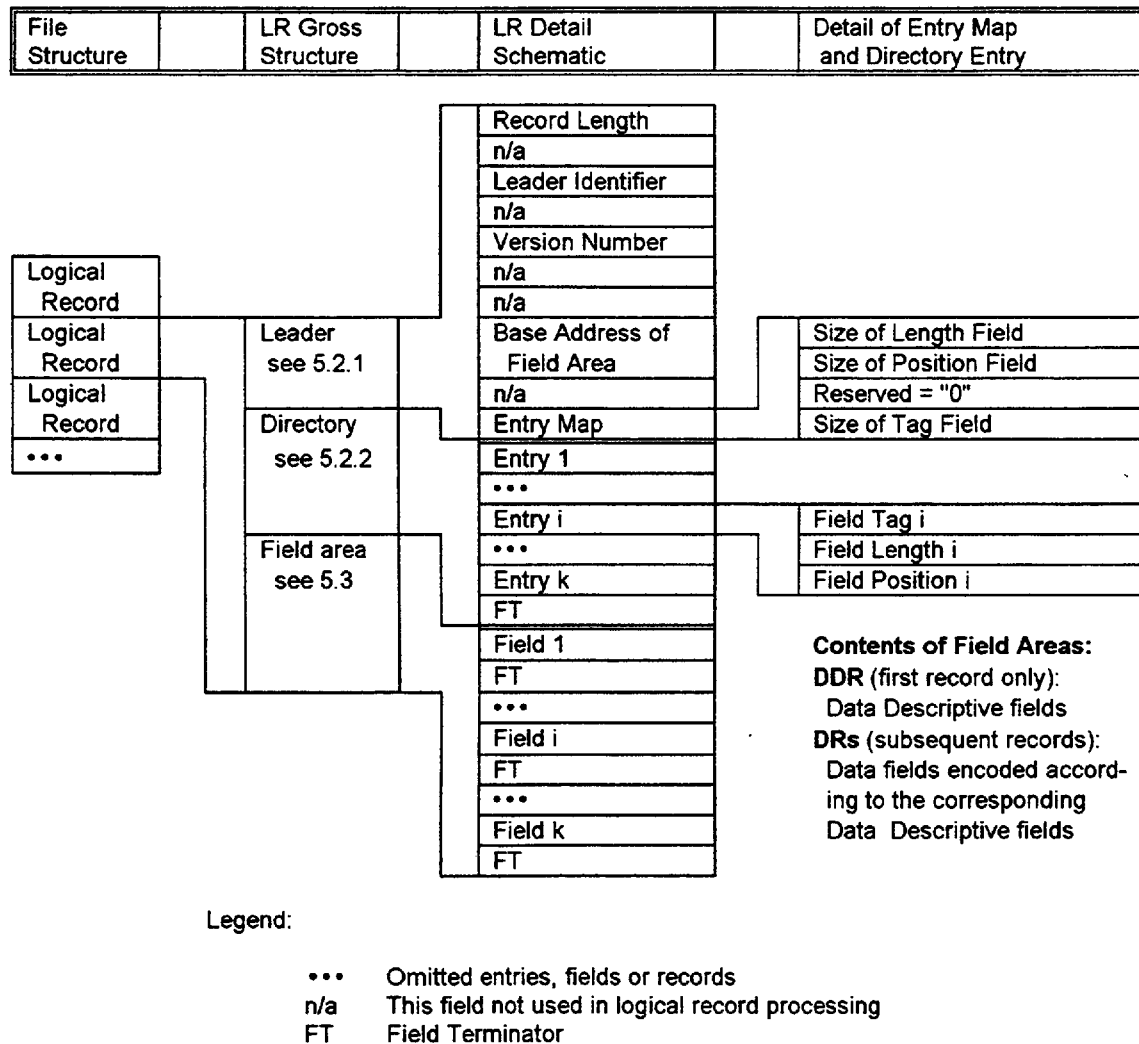


**Figure 1 - Schematic of ISO/IEC 8211 File and Logical Records**

The leaders of all records contain the parameters necessary to read records and disaggregate the directory into its entries. Additionally the DDR leader contains a few data descriptive parameters applicable to the entire file.

The directories of all records contain the parameters necessary to identify and locate each field in the Field Area.

In the first record only, the DDR, the Field Area contains data descriptive fields, each containing the information necessary to decode the user data in the Field Areas of the following DRs.

## 5.1 File and logical record structure

This International Standard specifies Data Descriptive Files (DDFs), each comprising Logical Records (LRs).

The files may be recorded on a physically dismountable medium or transported through a communications system. When recorded on a medium for which an International Standard for volume and file structure exists, each file shall have the required International Standard interchange file labels or headers for the particular medium. The file, when recorded on a suitable media volume, may be accessed sequentially or randomly.

An ISO/IEC 8211 interchange file set shall comprise one or more instances of the collection of items depicted in figure 2.

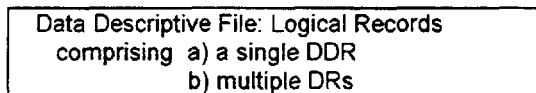


Figure 2 - File Schematic Representation

The DDR is the first logical record of each DDF.

### 5.1.1 Interchange logical records

The logical records specified in this International Standard shall be written, blocked and spanned, across the fixed-length, media physical records (blocks or sectors) or transmission packets without any further record demarcations.

### 5.1.2 Padding of records and media blocks

Any unused portion of a fixed-length, logical record without a leader/directory and any unused portion of a media block or its equivalent shall be padded with the CIRCUMFLEX character, (5/14).

NOTE 3 - This specification gives a satisfactory and detectable end of file condition under all known circumstances. The logic for processing the end-of-data and end-of-file condition is discussed in C.1.6.

## 5.2 Logical records - leaders and directories

This subclause contains specifications for logical records and their fields and for the parameters of leaders and directories of logical records which are involved in record processing. Both numeric and binary encodings of some leader and directory parameters pertaining to the octet counts are supported and are described in 5.2.1.6. The DDR leader parameters involved with data description are described in 6.1.

The logical records shall consist of the areas shown in figure 3.

| Name of Area | Length           |
|--------------|------------------|
| Leader       | 24               |
| Directory    | $k \times p + 1$ |
| Field Area   | variable         |

where  $k$  is the number of directory entries and  
 $p$  is the number of octets in a directory  
 entry of the record.

Figure 3 - Logical Record Schematic

### 5.2.1 Logical record leader

Each LR Leader shall consist of the fields shown in figure 4 and which are further specified in 5.2.1.1 to 5.2.1.5 and 6.1.

The "Use" column in figure 4 designates the use of each field as a record processing control. "All" indicates that a meaningful value exists in the field for all records, "DDR" indicates that a meaningful value exists in the field for the DDR only.

| RP | Len. | Use | Field Name                       | LR Leader       |           | Reference |
|----|------|-----|----------------------------------|-----------------|-----------|-----------|
|    |      |     |                                  | DDR             | DR        |           |
| 0  | 5    | All | Record Length                    | octets          | octets    | 5.2.1.1   |
| 5  | 1    | DDR | Interchange Level                | "1"   "2"   "3" | SPACE     | 6.1.1     |
| 6  | 1    | All | Leader Identifier                | "L"             | "D"   "R" | 5.2.1.2   |
| 7  | 1    | DDR | Inline Code Extension Indicator  | character       | SPACE     | 6.1.2     |
| 8  | 1    | DDR | Version number                   | SPACE   "1"     | SPACE     | 5.2.1.3   |
| 9  | 1    | DDR | Application Indicator            | character       | SPACE     | 6.1.3     |
| 10 | 2    | DDR | Field Control Length             | digits          | SPACEs    | 6.1.4     |
| 12 | 5    | All | Base Address of Field Area       | octets          | octets    | 5.2.1.4   |
| 17 | 3    | DDR | Extended Character Set Indicator | characters      | SPACEs    | 6.1.5     |
| 20 | 4    | All | Entry Map                        | digits          | digits    | 5.2.1.5   |

Figure 4 - LR Leader Schematic

#### 5.2.1.1 record length field (LR RP 0-4)

This field shall specify the total length of the LR in octets including the five octets of this field (see 5.2.1.6 for the syntax).

#### 5.2.1.2 leader identifier field (LR RP 6)

This field shall identify the type of logical record and specify the nature of the leader and directories.

When all remaining data records in a DDF would have identical leaders and directories, it is permissible to place an "R" in its Leader Identifier field and to omit the leader and directory from subsequent records,

The allowed values and meanings of the leader identifier shall be:

|   |                                                                                                                                                                                                                   |
|---|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| L | means that the record is the Data Descriptive Record                                                                                                                                                              |
| D | means that the record is a Data Record and that the next data record has a leader and directory.                                                                                                                  |
| R | means that the record is a Data Record and that a leader and directory will not be found in any of the subsequent LRs. The leader and directory of the current LR shall be applied to each of the subsequent LRs. |

#### 5.2.1.3 ISO/IEC 8211 version number (LR RP 8)

In the DDR, this field shall contain the identification of the version of ISO/IEC 8211 to which an interchange file conforms. The version number places conditions and constraints on both record and data description processing. The allowed values and meanings shall be:

| Value | Meaning             |
|-------|---------------------|
| SPACE | ISO/IEC 8211 - 1985 |
| "1"   | this version        |

This field is unused in the remaining logical records, the DRs, and shall contain SPACE.

#### 5.2.1.4 Base address of field area (LR RP 12-16)

This field shall specify the relative position of the first field of the Field Area and its value shall be equal to the combined length in octets of the leader and directory including the field terminator at the end of the directory (see 5.2.1.6).

**5.2.1.5 Entry map field (LR RP 20-23)**

| RP | Name of Subfield                    | Length | Contents   |
|----|-------------------------------------|--------|------------|
| 20 | Size of Field Length Field          | 1      | digit      |
| 21 | Size of Field Position Field        | 1      | digit      |
| 22 | Reserved for future standardization | 1      | DIGIT ZERO |
| 23 | Size of Field Tag Field             | 1      | digit      |

**Figure 5 - LR Entry Map Schematic**

This field specifies the lengths of directory entry subfields in each LR and shall consist of the subfields shown in figure 5 (see 5.2.1.6).

**5.2.1.5.1 Size of field length field (LR RP 20)**

This subfield shall specify the size in octets of the Field Length field of the directory entries and shall be a digit in the range of "1" to "9" inclusive.

**5.2.1.5.2 Size of field position field (LR RP 21)**

This subfield shall specify the size in octets of the Field Position field of the directory entries and shall be a digit in the range of "1" to "9" inclusive.

**5.2.1.5.3 Reserved for future standardization (LR RP 22)**

This subfield is reserved for future standardization as an extended entry map. Its value shall be the digit "0".

**5.2.1.5.4 Size of field tag field (LR RP 23)**

This subfield shall specify the size in octets of the Field Tag field of the directory entries and shall be a digit in the range of "1" to "7" inclusive. The value in this subfield shall be the same in all LRs.

**5.2.1.6 Alternate forms of counts and field positions**

This International Standard provides for both numeric and binary encoding of octet counts and positions.

The content of the fifth octet of the DDR Record Length (DDR RP 4) shall signal the use of the alternate forms. The allowed values shall have the following meanings:

| Value   | Signifying that:                                                                  |
|---------|-----------------------------------------------------------------------------------|
| "0"-"9" | the character form is used (the flag is the least significant digit of the value) |
| "B"     | the MSOF binary form is used                                                      |
| "b"     | the LSOF binary form is used                                                      |

The forms shall be mutually exclusive and the same form shall be used in the leader/directories throughout the file.

The leaders and directories of the LRs shall have one of the following forms for the contents of the Record Length field, the Base Address of Field Area field, the Field Length field and the Field Position field:

a) a character form comprising decimal digits, in which these fields shall contain a right-adjusted decimal integer filled on the left by ZEROs. A value of zero in the Record Length field shall specify a record length in excess of 99999.

b) a binary form comprising a signed binary integer and, in the DDR Record Length field only, a form flag. The first four octets of the Record Length field and Base Address of Field Area field shall be used to record a 32-bit signed binary integer. Each instance of an Entry Map shall be limited to a single set of values from the following list: 1,1,0,2; 2,2,0,2; 2,2,0,4 or 4,4,0,4. The values of the field length and field position shall be either 8-bit, 16-bit or 32-bit signed binary integers as appropriate depending upon the field sizes specified in the Entry Map. The maximum size of DDF record components shall be limited to the maximum value which can be placed in the corresponding control field. The unused positions, DR RP 4 and LR RP 16, shall contain a "0".

## 5.2.2 Logical record directory

The LR directory shall consist of repeated LR directory entries, the field lengths of which shall be specified in the entry map field (see 5.2.1.5). The LR directory shall contain one LR directory entry for each field in the Field Area and shall end with a field terminator.

The LR directory entry specifies the location and length of the corresponding field in the field area and shall consist of the subfields shown in figure 6. Each entry shall contain a Field Tag, Field Length, and Field Position in that sequence, and shall consist of  $m+n+t$  octets, where:

|     |                                         |
|-----|-----------------------------------------|
| $m$ | is the size of the Field Length Field   |
| $n$ | is the size of the Field Position Field |
| $t$ | is the size of the Field Tag Field      |

The LR directory entries shall be in one to one correspondence with the fields and in the same order.

The location, length and contents of each directory entry are specified in figure 6 and their meanings are specified in clauses 5.2.2.1 to 5.2.2.3.

| RP           | Name of Field  | Length | Contents     |
|--------------|----------------|--------|--------------|
| $p(i-1)$     | Field Tag      | $t$    | alphanumeric |
| $p(i-1)+t$   | Field Length   | $m$    | digits       |
| $p(i-1)+t+m$ | Field Position | $n$    | digits       |

where  $p = t + m + n$  and  $i$  = the index of the directory entry.

Figure 6 - LR Directory Entry Schematic

### 5.2.2.1 Field tag field

This field shall contain a Field Tag identifying the corresponding field and shall consist of between one and seven alphanumeric characters as specified in the Entry Map (see 5.2.1.5.4).

### 5.2.2.2 Field length field

This field specifies the length in octets of the corresponding field (see 5.2.1.6). The length of the field shall include the field terminator.

### 5.2.2.3 Field position field

This field specifies the relative position of the first octet in the corresponding field (see 5.2.1.6). It shall contain a value which when added to the base address of the Field Area shall result in the relative position of the first octet of the field referenced.

NOTE 4 - The field position of the first field following the directory is zero.

## 5.3 Logical record field areas

This subclause contains the general requirements of the field areas of logical records. Additional requirements for field areas related to the details of data description are found in 6.2, 6.3 and 6.4.

The field area of the first logical record (the DDR) shall contain Data Descriptive Fields describing the Data Fields of the file. The field areas of the subsequent logical records (the DRs) shall contain the Data Fields as described in the Data Descriptive Fields of the DDR. The field tags in the DDR and DR shall be used to associate the Data Descriptive Fields with the corresponding Data Fields.

The order of the tagged fields of a field area shall be the same as the order of the tags in the directory of the record.

### 5.3.1 Field area of the DDR

The Field Area of the DDR shall contain one and only one Data Descriptive Field for each data Field Tag that may be used in the Data Records.

### 5.3.2 The field areas of the DRs

The data fields of the Field Area of the DRs shall contain the actual data to be interchanged. Each data field shall be associated with an entry in the directory of the DR containing it, and the tag in that directory entry, by reference to the corresponding tag in the DDR and its Data Descriptive Field, may be used to interpret the data field.

A data field may be terminated prematurely at the end of a subfield prior to the end of its data description. The subfields shall contain the data elements corresponding to the expanded array descriptor contained in the corresponding DDR field. Each field shall end with the field terminator. In a delimited structure missing data elements shall be represented by successive delimiters. The null values to be associated with the missing data elements of a fixed-length subfield are not defined. In a user data field, a single terminal subfield delimiter or a terminal string of adjacent subfield delimiters may be omitted.

A data field, described in the DDR, may have zero or more instances in a DR Field Area. The directory entries corresponding to missing fields shall be omitted from the directory unless required to describe the data structure of a level 3 DDF (see 6.1.1).

#### 5.3.2.1 User data fields

Each data field shall conform to the Data Descriptive Field associated with the same Field Tag located in the DDR. Each data field shall end with a Field Terminator. The description of the user data is further defined in 6.3 and 6.4.

Character data fields whose termination is determined by a field length, format width, user determined delimiter or field terminator may contain a unit terminator character which shall be treated by an implementation as a user data character.

##### 5.3.2.1.1 Elementary data fields

The data fields of the DR shall contain a single string of octets terminated by the field terminator.

##### 5.3.2.1.2 Compound data fields

The data fields of the DR shall contain a single string of octets comprising subfields, terminated by the field terminator.

## 6 Description of user data types and structures

This clause contains specifications for the description of the user's data fields. See clause C.2 and annex D for an informal discussion of the constructs of this clause.

### 6.1 DDR leader fields related to data description

This subclause contains additional specifications for the DDR leader which are related to data description and which are applicable to the entire file.

In addition to the fields contained in all LRs, the DDR leader shall also contain meaningful values in the DDR specific fields, as shown in figure 4 and further specified in 6.1.1 to 6.1.5.

#### 6.1.1 Interchange level field (DDR RP 5)

This field shall specify the level of the interchange file. The content of this field shall be the digit "1", "2" or "3" and these shall have the following meanings:

|     |                            |
|-----|----------------------------|
| "1" | the file is a level 1 file |
| "2" | the file is a level 2 file |
| "3" | the file is a level 3 file |

In a level 1 file the Data Fields of the DDR shall contain octet strings which are not differentiated into ISO/IEC 8211 subfields.

In level 2 and level 3 files the Data Fields of the DDR shall contain compound data fields which are described by the corresponding Data Descriptive Fields. In a level 3 file there is an additional inter-field hierarchical structure.

For level 3 files, the tags and fields of the LR shall be in the same order as the preorder traversal sequence of the corresponding fields of the ordered, rooted tree of the LR. In the absence of the 0...1 Field Tag, the first Field Tag of the preorder traversal sequence shall be the root node of the tree.

NOTE 5 - The clauses which have substantive reference to level specifications are:

| Level | Relevant clauses    |
|-------|---------------------|
| 1     | 6.3                 |
| 2     | 6.4                 |
| 3     | 6.2.1.3, 6.2.6, 6.4 |

### 6.1.2 Inline code extension indicator (DDR RP 7)

This field shall specify the coded character set methodology used and if additional inline coded character sets are used. SPACE shall mean that there are no character sets other than the basic character set used in the file.

NOTE 6 - The other values for this field are specified in clause 7.

### 6.1.3 Application indicator field (DDR RP 9)

This field contains identification of the user application standard under which a file was produced. SPACE shall mean that there is no identification of the application. The value "A" shall specify a reference to another standard (see 6.1.3.1). The values "B" to "Z" inclusive are reserved for future standardization. Other values are reserved for private use.

#### 6.1.3.1 Reference to other standards

The value "A" shall mean that another standard is used to specify the content and semantics of the transferred information and the numerical form of the Object Identifier Arcs of the applicable standard will be found in the first subfield of the User Application field (tag = 0...2, see 6.2.3) followed by an UT when additional user text is required.

NOTE 7 - The syntax and semantics of the Object Identifier Arcs are specified in ISO/IEC 8824-1 Annexes B, C and D.

### 6.1.4 Field control length field (DDR RP 10-11)

This field shall specify the number of octets of the Data Descriptive Field devoted to data element type and structure codes, delimiters, truncated escape sequence (if any) and other positions reserved for future standardization.

The contents of this field shall be:

| Level  | Non-extended sets | Extended sets |
|--------|-------------------|---------------|
| 1      | "00"              | "03"          |
| 2 or 3 | "06"              | "09"          |

where non-extended and extended sets refer to the use of fieldwise code extensions facilities (see 6.1.5).

### 6.1.5 Extended character set indicator field (DDR RP 17-19)

This field shall specify the filewise default coded character set used or if fieldwise default character sets are defined. SPACES shall mean that there are no default character sets defined for this file.

NOTE 8 - The values for this field are specified in clause 7.

## 6.2 Special field tags (tags = 0...0 to 0...9)

This subclause contains requirements for special field tags and the contents of their Data Descriptive Fields which are applicable to the entire file and which are not applicable to an isolated data field.

### 6.2.1 File control field (tag = 0...0)

The file control field shall be present only in the DDR, shall have the tag 0...0 and shall contain the following subfields (see figure 7):

- a) the field controls (if any),
- b) an optional external file title,

- c) for level 3 only, a unit terminator followed by a list of field tag pairs,  
d) terminated by a field terminator.

|                |                     |    |                         |    |
|----------------|---------------------|----|-------------------------|----|
| Field Controls | External File Title | UT | List of Field Tag Pairs | FT |
|----------------|---------------------|----|-------------------------|----|

**Figure 7 - File Control Field Schematic**

#### 6.2.1.1 Field control field

These field controls are a special case of the field controls described in 6.3 and 6.4.2. The first four octets for level 2 and 3 files shall contain the value "0000".

NOTE 9 - The printable graphics and escape sequence of the field controls are applicable only to the External File Title as there are no corresponding DR fields.

#### 6.2.1.2 External file title field

This field shall specify an optional external file title. If present it shall contain a character string which shall be an external descriptive name for the interchange file.

#### 6.2.1.3 List of Field tag pairs

In a level 3 DDF, many ordered rooted trees may be described by the use of the list of field tag pairs and the preorder traversal sequence of the tree. See annex F for technical background.

NOTE 10 - Trees which cannot be represented by this method can be represented by the method of 6.2.6.

The list of field tag pairs is the parent/offspring binary relation of the tags of the DDR which with the preorder traversal sequence of the DDR describe a generic tree structure for the DDF. These pairs may be placed in the list in any sequence and shall be logically contiguous. Field tags 0...0 and 0...2 to 0...9 inclusive shall not participate in the structure specification.

NOTE 11 - The variable hierarchical data structures permitted in the DR are described by supplying the preorder traversal sequences of the data tree and a list of the field tag pairs expressing the parent/offspring logical association between the nodes of the data tree. The data structures of the DR shall be derived from the generic data tree by the repetition or deletion of nodes and subtrees. The preorder traversal sequence of the generic data tree is supplied by the order of the DDR directory entries. The preorder traversal sequence of each instance of a derived data tree in a DR is supplied by the sequence of directory entries in the DR. See annex F for further explanation of ordered field tag pairs. A forest may be represented as an ordered, rooted tree by creating ordered links from the root node of the first tree to the root nodes of the remaining trees in the proper order.

#### 6.2.2 Record identifier field (tag = 0...1)

The optional Record Identifier Field, if present, is a normal user Data Field, whose content is specified by the user. Each DR shall contain zero or one Record Identifier Fields. The content of the Record Identifier Field shall conform to the corresponding DDR Data Descriptive Field.

If used, the field corresponding to field tag 0...1 shall be the Record Identifier Field, and except for the LINKS field (tag = 0...9, see 6.2.6), shall occur first in the DR directory and shall be the root of any hierarchy (see 6.2.1.3).

#### 6.2.3 User application field (tag 0...2)

If used, the field with the field tag 0...2 shall be present only in the DDR and shall contain application specific information. An implementation shall pass this field to the user for processing.

NOTE 12 - The contents of this field are determined by the user and may be used to convey any augmented file description (e.g., file attributes relevant to the interchange) or ancillary file processing controls or application information. See 6.1.3.1 for the use of this field to identify other standards.

#### 6.2.4 Announcer sequence or feature identifier field (tag 0...3)

If used, the field with the tag 0...3 shall be present only in the DDR and shall contain the set of announcer sequences or feature identifiers for extended code set services, see clause 7.2.4 and 7.3.3. Any control characters in this field shall be encoded as their ISO/IEC 6429 mnemonics (see B.2.6). An implementation shall pass this field to the user for processing.

#### 6.2.5 Fields reserved for future standardization

Fields with tags 0...4 through 0...8 shall be reserved for future standardization and shall not be present in the DDR.

### 6.2.6 Recursive tree LINKS field (tag = 0...9)

NOTE 13 - This field is necessary to describe recursive trees which cannot be described by the list of Field Tag Pairs.

In each DR, a recursive tree shall be described by the left- and right-hand links of the corresponding binary tree. These links shall be placed in a reserved field having the field tag 0...9, and the following data description (see 6.4):

2f00;&LINKS&\*LLINK!RLINK&(intfmt); where

|               |                                       |                    |
|---------------|---------------------------------------|--------------------|
| <i>intfmt</i> | is an integer format control,         | = 1, l(d) or B2w,  |
| <i>t</i>      | is the data type code;                | = 1 for l formats, |
|               |                                       | = 5 for B formats, |
| <i>d</i>      | is the subfield width of an l format, |                    |
| <i>w</i>      | is the precision of a B format.       |                    |

When present, the LINKS field shall be the first field in the DR directory and field area and shall not participate in the tree structure. The remaining DR field tags and fields shall occur in the preorder traversal sequence of the tree. The left- and right-hand links to the offspring fields shall occur in the preorder traversal sequence and shall comprise the indices of the directory entries of the corresponding fields (see the example of F.2). When there is no link, the corresponding left- and right-link shall be zero. When present, the tree structure defined by the LINKS field shall supersede that defined by the field tag pairs.

The LINKS field may be used to define non-recursive trees.

NOTE 14 - A recursive tree cannot occur in the DDR and the DDR tree is described by its field tag pairs and preorder traversal sequence. The preorder traversal sequence of the corresponding binary tree is the same as that of the ordered, rooted tree.

### 6.2.7 Order of special field tags in the DDR

When present, field tags 0...0 to 0...9 inclusive shall occur first in the DDR directory and in ascending numeric order.

## 6.3 Data descriptive fields in level 1 files

This subclause specifies the content of the Data Descriptive Fields of level 1 files. The user data of the corresponding Data Fields comprises octet strings without division into subfields.

A data descriptive field of a level 1 file shall comprise the following:

|                |                 |    |
|----------------|-----------------|----|
| Field Controls | Data Field Name | FT |
|----------------|-----------------|----|

A level 1 interchange file may have one Field Control parameter (the truncated escape sequence) whose presence is conditional upon the use of fieldwise default extended coded character sets. Its use and values are specified in clause 7.

The Data Descriptive Fields of level 1 files shall contain an optional Data Field Name which serves to identify the data field. The Data Field Name may be any string of characters.

## 6.4 Data descriptive fields in level 2 and 3 files

This subclause specifies the content of the Data Descriptive Fields of level 2 and 3 files. The user data of the corresponding Data Fields comprises subfields.

The data structures specified in this subclause shall be elementary, vector and array structures containing character strings, implicit point, explicit point, explicit point scaled, character mode bit string, bit string, binary forms and mixed data types. The data descriptive field shall contain control information, printable graphics, field name, array descriptors and data field format information as specified in 6.4.1 to 6.4.4 and clause 7.

### 6.4.1 Tabular summary of data descriptive fields of level 2 and 3 files

The data descriptive fields for compound data fields shall comprise the Field Control field, the Data Field Name field, Array descriptors and Format Controls. A schematic of a Data Descriptive Field is given in figure 8. Missing elements of the data description shall be represented by adjacent delimiters and a terminal string of adjacent delimiters shall not be replaced by the field delimiter. Table 1 specifies the use of delimiters in Data Descriptive Fields and in Data Fields. Table 2 summarizes the data descriptive field components which are specified in this clause and defined further in 6.4.2 to 6.4.4.

|                |                 |    |                  |    |                 |    |
|----------------|-----------------|----|------------------|----|-----------------|----|
| Field Controls | Data Field Name | UT | Array descriptor | UT | Format Controls | FT |
|----------------|-----------------|----|------------------|----|-----------------|----|

Figure 8 - Schematic of Level 2 and 3 Data Descriptive Fields

Table 1 - Delimiters and Their Uses

| Delimiter | Print Symbol | Record                  | Usage                                                                                                                                                                                               |
|-----------|--------------|-------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| FT [1]    | ; [2]        | DDR, DR                 | Field terminator                                                                                                                                                                                    |
| UT [1]    | & [2]        | DR<br>DDR<br>DDR<br>DDR | Unit terminator<br>a) To delimit subfields in a field not specified by a format<br>b) To delimit the Data Field Name<br>c) To delimit the External File Title<br>d) To delimit the array descriptor |
| "!"       | !            | DDR                     | To delimit subfield labels within a vector label                                                                                                                                                    |
| "*"       | *            | DDR                     | To delimit the vector labels within a Cartesian label.                                                                                                                                              |
| "\""      | \\           | DDR                     | To delimit the array descriptors of concatenated structures.                                                                                                                                        |

[1] These characters are defined by the BCS codes, IS2 and IS1 ((1/14) and (1/15)). With the use of ISO/IEC 10646 codes they are represented by the appropriate multi-octet code (see clause 7 and ISO/IEC 10646 clause 16).

[2] The print symbols, ";", "&" and "\" are the graphic characters used in examples throughout this document to represent the standard delimiters, FT and UT, which are non-printing characters in many systems. See 6.4.2.4.

Table 2 - Data Descriptive Field Components

| Data Structure | RP 0, 1 | Data Field Name & Array descriptor & Format Controls;   |
|----------------|---------|---------------------------------------------------------|
| Elementary     | 0d      | ['Name'] & ['subfield label'] & ['fctrl'];              |
| Vector         | 1d      | ['Name'] & ['Vector label'] & ['fctrl'];                |
| Array          | 2d      | ['Name'] & ['Array descriptor'] & ['fctrl'];            |
| Concatenated   | 3d      | ['Name'] & 'Concatenated data description' & ['fctrl']; |

The syntax of table 2 is defined below:

- a) RP 0 contains the data structure code (see 6.4.2.1), RP 1 = "d" is a data type code (see 6.4.2.2).
- b) 'Name' is a data field name (see 6.4.3.1).
- c) 'Array descriptor' specifies one of five forms which identifies subfields and structure. These forms are: concatenated structure, numeric array descriptor, Cartesian label, vector label or subfield label (see 6.4.3.2).
- d) 'fctrl' specifies a format control (see 6.4.3.3).
- e) [ ] specifies the optional or conditional presence of the item.

#### 6.4.2 Field controls

This clause specifies the position, values and meaning of the field control parameters which describe the contents of the DR user data fields. The length of the Field Controls field is specified by the contents of DDR RP 10-11. Level 2 and 3 interchange files have the following field control parameters.