

INTERNATIONAL  
STANDARD

**ISO/IEC**  
**8613-1**

Second edition  
1994-12-15

---

---

**Information technology — Open  
Document Architecture (ODA) and  
Interchange Format: Introduction and  
general principles**

*Technologies de l'information — Architecture des documents ouverts  
(ODA) et format d'échange: Introduction et principes généraux*



Reference number  
ISO/IEC 8613-1:1994(E)  
**Documents**  
AVA Document Sharing Hub

## CONTENTS

|                                                                                            | <i>Page</i> |
|--------------------------------------------------------------------------------------------|-------------|
| 1 Scope .....                                                                              | 1           |
| 2 Normative references .....                                                               | 1           |
| 2.1 Identical Recommendations   International Standards .....                              | 2           |
| 2.2 Paired Recommendations   International Standards equivalent in technical content ..... | 2           |
| 2.3 Additional references .....                                                            | 2           |
| 3 Definitions .....                                                                        | 3           |
| 4 Abbreviations .....                                                                      | 12          |
| 5 Conventions .....                                                                        | 13          |
| 5.1 Names of attributes and parameters .....                                               | 13          |
| 5.2 Names of attribute and parameter values .....                                          | 13          |
| 5.3 Component and component description .....                                              | 13          |
| 5.4 Introduction of terms .....                                                            | 13          |
| 5.5 Representation of degrees .....                                                        | 14          |
| 6 General concept of ODA .....                                                             | 14          |
| 6.1 Purpose of ODA .....                                                                   | 14          |
| 6.2 Overall concept of ODA .....                                                           | 14          |
| 6.2.1 Logical structure and layout structure .....                                         | 14          |
| 6.2.2 Content portions .....                                                               | 16          |
| 6.2.3 Content architectures .....                                                          | 16          |
| 6.2.4 Attributes .....                                                                     | 17          |
| 6.2.5 Relations between logical structure and layout structure .....                       | 17          |
| 6.2.6 Specific and generic structures .....                                                | 17          |
| 6.2.7 Document profile .....                                                               | 19          |
| 6.2.8 Generic-document .....                                                               | 19          |
| 6.2.9 Protected parts .....                                                                | 19          |
| 6.3 Document processing model .....                                                        | 19          |
| 6.3.1 Relationships of ODA to document processing .....                                    | 19          |
| 6.3.2 Editing process .....                                                                | 19          |
| 6.3.3 Layout process .....                                                                 | 19          |
| 6.3.4 Imaging process .....                                                                | 20          |

© ISO/IEC 1994

All rights reserved. Unless otherwise specified, no part of this publication may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm, without permission in writing from the publisher.

ISO/IEC Copyright Office • Case postale 56 • CH-1211 Genève 20 • Switzerland

Printed in Switzerland

- 7 Overview of the ITU-T Rec. T.410-Series | ISO/IEC 8613 ..... 20
  - 7.1 ITU-T Rec. T.411 | ISO/IEC 8613-1: Introduction and general principles ..... 20
  - 7.2 ITU-T Rec. T.412 | ISO/IEC 8613-2: Document structures ..... 21
  - 7.3 ITU-T Rec. T.414 | ISO/IEC 8613-4: Document profile ..... 21
  - 7.4 ITU-T Rec. T.415 | ISO/IEC 8613-5: Open document interchange format..... 21
  - 7.5 ITU-T Rec. T.416 | ISO/IEC 8613-6: Character content architectures..... 21
  - 7.6 ITU-T Rec. T.417 | ISO/IEC 8613-7: Raster graphics content architectures..... 21
  - 7.7 ITU-T Rec. T.418 | ISO/IEC 8613-8: Geometric graphics content architectures ..... 22
  - 7.8 ISO/IEC 8613-10: Formal specifications..... 22
- 8 Inter-dependencies of the Specifications..... 22
- 9 Conformance ..... 23
- 10 Document application profile..... 23
  - 10.1 General principles for defining a document application profile ..... 24
  - 10.2 Rules for defining a document application profile..... 24
    - 10.2.1 Rules for defining a document architecture class ..... 24
    - 10.2.2 Rules for defining a content architecture class ..... 25
    - 10.2.3 Rules for defining a document profile ..... 27
    - 10.2.4 Rules for selecting the interchange format class..... 27

LIST OF FIGURES

|                                                                               | <i>Page</i> |
|-------------------------------------------------------------------------------|-------------|
| Figure 1 – Example of a logical view of a document .....                      | 15          |
| Figure 2 – Possible correspondence between logical and layout objects .....   | 16          |
| Figure 3 – Example of a page layout.....                                      | 18          |
| Figure 4 – Basic document processing model.....                               | 20          |
| Figure G.1 – Data stream conformance to ITU-T Rec. T.410   ISO/IEC 8613 ..... | 76          |

LIST OF TABLES

|                                                            |    |
|------------------------------------------------------------|----|
| Table F.1 – List of attributes and data types.....         | 57 |
| Table F.2 – List of control functions and data types ..... | 74 |

## 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 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 8613-1 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 18, *Document processing and related communication*, in collaboration with ITU-T. The identical text is published as ITU-T Rec. T.411.

This second edition cancels and replaces the first edition (ISO 8613-1:1989), which has been technically revised.

ISO/IEC 8613 consists of the following parts, under the general title *Information technology — Open Document Architecture (ODA) and Interchange Format*:

- *Part 1: Introduction and general principles*
- *Part 2: Document structures*
- *Part 3: Abstract interface for the manipulation of ODA documents*
- *Part 4: Document profile*
- *Part 5: Open Document Interchange Format*
- *Part 6: Character content architectures*
- *Part 7: Raster graphics content architectures*
- *Part 8: Geometric graphics content architectures*
- *Part 9: Audio content architectures*
- *Part 10: Formal specifications*
- *Part 11: Tabular structures and tabular layout*
- *Part 12: Identification of document fragments*
- *Part 13: Spreadsheet*
- *Part 14: Temporal relationships and non-linear structures*

Annexes C, D and E form an integral part of this part of ISO/IEC 8613. Annexes A, B, F and G are for information only.

## Introduction

This ITU-T Recommendation | International Standard was prepared as a joint publication by TSS Study Group 8 and ISO/IEC Joint Technical Committee 1.

At present, the ITU-T series of Recommendations in the T.410-Series | International Standard ISO/IEC 8613 consists of:

- Introduction and general principles;
- Document structures;
- Document profile;
- Open document interchange format;
- Character content architectures;
- Raster graphics content architectures;
- Geometric graphics content architectures;
- Formal specifications.

(The formal specification is applicable to ISO/IEC 8613 only.)

Further Recommendations | International Standards may be added to this series of ITU-T Recommendations | International Standard.

Development of this series of ITU-T Recommendations | International Standards was originally in parallel with the ECMA 101 standard: *Open Document Architecture*.

This series of ITU-T Recommendations | International Standards is a new edition of the CCITT T.410-Series of Recommendations (1988) and ISO/IEC 8613 (1989).

Significant technical changes are the inclusion of the following amendments as agreed by TSS and ISO/IEC:

- Alternative representation
- Annex on use of MHS/MOTIS
- Colour
- Conformance Testing annex
- Document Application Profile Proforma and Notation
- Security
- Streams
- Styles
- Tiled raster graphics

In addition, a number of technical corrigenda have been applied to this series of ITU-T Recommendations | International Standard.

This ITU-T Recommendation | International Standard contains eight annexes:

- Annex A (non-integral) gives references to other standards and registers;
- Annex B (non-integral) gives the relationship of this series of ITU-T Recommendations | International Standards with other CCITT Recommendations and International Standards;
- Annex C (integral) specifies the method of recording documents in conformance with this series of ITU-T Recommendations | International Standards on volume and file structures of interchangeable storage media in conformance with international standards for volume and file structures;
- Annex D (integral) gives the principles for the assignment of ASN.1 object identifier values for this series of ITU-T Recommendations | International Standards;
- Annex E (integral) describes the use of MHS to interchange documents conforming to this series of ITU-T Recommendations | International Standards;
- Annex F (non-integral) defines a standardized proforma for the specification of document application profiles based on this series of ITU-T Recommendations | International Standards;
- Annex G (non-integral) gives directions for the development of a conformance testing methodology for this series of ITU-T Recommendations | International Standards.



## INTERNATIONAL STANDARD

## ITU-T RECOMMENDATION

# INFORMATION TECHNOLOGY – OPEN DOCUMENT ARCHITECTURE (ODA) AND INTERCHANGE FORMAT: INTRODUCTION AND GENERAL PRINCIPLES

## 1 Scope

The purpose of the ITU-T Rec. T.410-Series | ISO/IEC 8613 is to facilitate the interchange of documents.

In the context of these Recommendations | International Standards, documents are to be items such as memoranda, letters, invoices, forms and reports, which may include pictures and tabular material. The content elements used within the documents may include graphic characters, raster graphics elements, and geometric graphics elements, all potentially within one document.

NOTE – These Recommendations | International Standards are designed to allow for extensions, including hypermedia features, spreadsheets and additional types of content such as audio and video.

In addition to content types defined in these Recommendations | International Standards, ODA also provides for arbitrary content types to be included in documents.

These Recommendations | International Standards apply to the interchange of documents by means of data communications or the exchange of storage media.

These Recommendations | International Standards provide for the interchange of documents for either or both of the following purposes:

- to allow presentation as intended by the originator;
- to allow processing, such as editing and reformatting.

The composition of a document in interchange can take several forms:

- formatted form, allowing presentation of the document;
- processable form, allowing processing of the document;
- formatted processable form, allowing both presentation and processing of the document.

These Recommendations | International Standards also provide for the interchange of ODA information structures used for the processing of interchanged documents.

This ITU-T Recommendation | International Standard:

- introduces this ITU-T T.410-Series of Recommendations | ISO/IEC 8613 as a whole;
- gives the references necessary for the ITU-T T.410-Series of Recommendations | ISO/IEC 8613;
- defines terms used in the context of the ITU-T T.410-Series of Recommendations | ISO/IEC 8613;
- presents the concepts of the document architecture;
- gives an overview of the ITU-T T.410-Series of Recommendations | ISO/IEC 8613;
- describes the inter-dependencies of the ITU-T T.410-Series of Recommendations | ISO/IEC 8613;
- defines conformance to the ITU-T T.410-Series of Recommendations | ISO/IEC 8613;
- gives rules for defining document application profiles.

## 2 Normative references

The following ITU-T/CCITT Recommendations and International Standards contain provisions which, through reference in this text, constitute provisions of this Recommendation | International Standard. At the time of publication, the editions indicated were valid. All Recommendations and Standards are subject to revision, and parties to agreements based on this Recommendation | International Standard are encouraged to investigate the possibility of applying the most recent edition of the Recommendations and Standards listed below. Members of IEC and ISO maintain registers of currently valid International Standards. The ITU-T Secretariat maintains a list of the currently valid ITU-T/CCITT Recommendations.

## 2.1 Identical Recommendations | International Standards

- ITU-T Recommendation T.412 (1993) | ISO/IEC 8613-2:1994, *Information technology – Open Document Architecture (ODA) and Interchange Format: Document structures.*
- ITU-T Recommendation T.414 (1993) | ISO/IEC 8613-4:1994, *Information technology – Open Document Architecture (ODA) and Interchange Format: Document profile.*
- ITU-T Recommendation T.415 (1993) | ISO/IEC 8613-5:1994, *Information technology – Open Document Architecture (ODA) and Interchange Format: Open Document Interchange Format.*
- ITU-T Recommendation T.416 (1993) | ISO/IEC 8613-6:1994, *Information technology – Open Document Architecture (ODA) and Interchange Format: Character content architectures.*
- ITU-T Recommendation T.417 (1993) | ISO/IEC 8613-7:1994, *Information technology – Open Document Architecture (ODA) and Interchange Format: Raster graphics content architectures.*
- ITU-T Recommendation T.418 (1993) | ISO/IEC 8613-8:1994, *Information technology – Open Document Architecture (ODA) and Interchange Format: Geometric graphics content architectures.*

## 2.2 Paired Recommendations | International Standards equivalent in technical content

- CCITT Recommendation X.208 (1988), *Specification of Abstract Syntax Notation One (ASN.1).*  
ISO/IEC 8824:1990, *Information technology – Open Systems Interconnection – Specification of Abstract Syntax Notation One (ASN.1).*
- CCITT Recommendation X.209 (1988), *Specification of basic encoding rules for Abstract Syntax Notation One (ASN.1).*  
ISO/IEC 8825:1990, *Information technology – Open Systems Interconnection – Specification of Basic Encoding Rules for Abstract Syntax Notation One (ASN.1).*
- CCITT Recommendation X.411 (1992), *Message Handling Systems: Message Transfer System: Abstract Service Definition and Procedures.*  
ISO/IEC 10021-4:1990, *Information technology – Text Communication – Message Oriented Text Interchange Systems (MOTIS) – Part 4: Message Transfer System: Abstract Service Definition and Procedures.*
- CCITT Recommendation X.420 (1992), *Message Handling Systems: Message Transfer System: Interpersonal Messaging System.*  
ISO/IEC 10021-7:1990, *Information technology – Text Communication – Message Oriented Text Interchange Systems (MOTIS) – Part 7: Interpersonal Messaging System.*

## 2.3 Additional references

- ISO 2022:1986, *Information processing – ISO 7-bit and 8-bit coded character sets – Code extension techniques.*
- ISO/IEC 8613-10:1991, *Information processing – Text and office systems – Office Document Architecture (ODA) and Interchange Format – Part 10: Formal specifications.*
- ISO/IEC 8632-1:1992, *Information technology – Computer graphics – Metafile for the storage and transfer of picture description information – Part 1: Functional specification.*
- ISO/IEC 8632-3:1992, *Information technology – Computer graphics – Metafile for the storage and transfer of picture description information – Part 3: Binary encoding.*
- ISO 8879:1986, *Information processing – Text and office systems – Standard Generalized Markup Language (SGML).*
- ISO 9069:1988, *Information processing – SGML support facilities – SGML Document Interchange Format (SDIF).*
- ISO 9293:1987, *Information processing – Volume and file structure of flexible disk cartridges for information interchange.*
- ISO/IEC 9541-1:1991, *Information technology – Font information interchange – Part 1: Architecture.*
- ISO 9660:1988, *Information processing – Volume and file structure of CD-ROM for information interchange.*

### 3 Definitions

For the purposes of this Recommendation | International Standard, the following definitions apply.

- 3.1 active position:** The point at which the action specified by the next character is to be effected.
- 3.2 aligned around:** A tabulation alignment that positions the sequence of character images for a specified character string such that the position point of the character image of the first instance of a specified group of characters within that string is positioned at the tabulation stop.
- 3.3 alternative description:** A description that represents a basic object that is intended to be used by the recipient in lieu of the primary description of that basic object when the primary description cannot be processed.
- 3.4 alternative subtree:** An alternative basic object description in conjunction with its associated content portion descriptions.
- 3.5 aspect ratio:** The ratio of the dimension of a pel array in the direction of the pel path to the dimension in the direction of the line progression.
- 3.6 assured reproduction area:** The rectangular area that remains on the nominal page after deducting an agreed allowance for edge losses.
- 3.7 attribute:** An element of a constituent of a document that has a name and a value and that expresses a characteristic of this constituent or a relationship with one or more constituents.
- 3.8 authenticity:** The property that the claimed data source can be verified to the satisfaction of the recipient.
- 3.9 available area:** The area determined by the document layout process into which the content portion is formatted by the content layout process.
- 3.10 basic component:** A basic logical or layout object, or an object class from which basic logical or layout objects may be derived.
- 3.11 basic layout object:** An object in the specific layout structure that has no subordinate.
- 3.12 basic logical object:** An object in the specific logical structure that has no subordinate.
- 3.13 basic measurement unit; BMU (abbreviation):** A unit of linear measurement equal to 1/1200 of 25,4 mm.  
NOTE – A locally defined scaling factor may be used to map the document to a particular imaging device.
- 3.14 basic value:** An attribute value, a control function parameter value or the value of any other capability that is unconditionally allowed in document interchange in the context of a given document application profile.
- 3.15 binding:** A pair comprising an identifier and a value, where the value may be of any type, may be specified by an expression, and is accessed through use of the binding identifier.
- 3.16 block:** A basic layout component that corresponds to a rectangular area within a frame or a page.
- 3.17 bottom edge:** The edge of the positioning area of a basic layout object that is in the direction of the line progression.
- 3.18 bottom left corner:** The corner of a layout object that is least progressed in the horizontal direction and most progressed in the vertical direction of a layout object.
- 3.19 bottom right corner:** The corner of a layout object that is most progressed both in the horizontal and vertical directions of a layout object.
- 3.20 centred:**
- a) The result of a layout or imaging process that positions the sequence of character images for a line such that the distance from the line home position to the position point of the first character image is approximately equal to the distance from the escapement point of the last character image to the end edge of the positioning area.
  - b) A tabulation alignment that positions the sequence of character images for a specified character string such that the distance from the position point of the first character image to the tabulation stop is approximately equal to the distance from the tabulation stop to the escapement point of the last character image.
- NOTE – The term *centred* is also used as a value in the sub-parameter “alignment” of the parameter “position” and in the attribute “block alignment”.
- 3.21 character:** A member of a set of elements used for the organization, control and representation of information.

- 3.22 character base line:** A line across a character image, in the horizontal direction when the character image is in its intended viewing orientation.
- 3.23 character image:** The human perceptible rendering of a character on a presentation medium.  
NOTE – The term *glyph* in ISO/IEC 9541 is equivalent to the term *character image* in this Specification.
- 3.24 character orientation:** The direction of the character base line relative to the characters.
- 3.25 character path:** The direction of progression of successive character images within a line box.
- 3.26 character sequence:** A sequence of characters intended to be presented as one or more lines.
- 3.27 character spacing (for constant spacing fonts only):** The distance between the position points of successive character images when the inter-character space equals zero.
- 3.28 clipped pel array:** The actual pel array to be imaged as determined by taking account of all clipping parameters.
- 3.29 colour component:** One of the dimensions of a colour space.
- 3.30 colour gamut:** The range of colours a given system is capable of reproducing.
- 3.31 complete generator set:** A constituent of a document consisting of a document root object class description and at least one level of subordinate object class descriptions which are used to control the creation and/or modification of the set of object descriptions representing a corresponding specific structure.
- 3.32 component:** An object or an object class.
- 3.33 composite component:** A composite logical or layout object, or an object class from which composite logical or layout objects may be derived.
- 3.34 composite layout object:** An object in the layout structure that has one or more subordinate objects.
- 3.35 composite logical object:** An object in the logical structure that has one or more subordinate objects.
- 3.36 confidentiality:** The property that information is not made available or disclosed to unauthorized individuals, entities or processes.  
NOTE – This property is limited here to preventing unauthorized semantic knowledge of a document or specified parts of it.
- 3.37 constant spacing:** The characteristic of a font wherein the distance between the position point and the escapement point is the same for all character images.
- 3.38 constituent:** A set of attributes that is one of the following types: a document profile, an object description, an object class description, a presentation style, a layout style, a content portion description or a protected part description.
- 3.39 content:** The information conveyed by the document, other than the structural information, and that is intended for human perception.
- 3.40 content architecture:** Rules for defining the internal structure and representation of the content of basic components in terms of a set of content elements, attributes and control functions, and guidelines for the presentation of the content.
- 3.41 content architecture class:** The rules for defining the internal structure and representation of the content of basic components in one of a set of forms defined for each type of content element.  
NOTE – Examples of content architecture classes are formatted form, processable form and formatted processable form in the case of character content elements.
- 3.42 content editing process:** The process that creates new content or modifies previous content.
- 3.43 content element:** A basic element of the content of a document.
- 3.44 content layout process:** The process that, interacting with the document layout process, consists of the formatting of content portions into available areas and the determination of the sizes of blocks in accordance with information contained in the presentation styles.
- 3.45 content portion:** The result of partitioning the content of a document according to its logical and/or layout structure.
- 3.46 content portion description:** A constituent of a document, representing a content portion that consists of content information and attributes to specify the properties of its content information.

**3.47 content type:** A category of content elements such as graphic characters, raster graphic elements and geometric graphic elements.

**3.48 control function:** An element of a character set that affects the recording, processing, transmission or interpretation of data, and that has a coded representation consisting of one or more bit combinations.

NOTE – Examples of control functions are Selected Graphic Rendition (SGR) in character content architectures and Set Line Type in geometric graphics content architectures.

**3.49 current layout position:** The identification of a lowest level frame which is maintained during the layout process for each layout stream which occurs.

**3.50 Cyan Magenta Yellow (Black); CMY(K) (abbreviation) colour space:** Colour space based on the subtractive colour mixture of Cyan (C), Magenta (M) and Yellow (Y) primaries with the optional inclusion of Black (K) as indicated by the parentheses.

**3.51 DAP-n data stream:** An ODA data stream in which the data elements are in accordance with a particular document application profile named *n*, defined in accordance with this Specification.

**3.52 data integrity:** The property that data has not been altered or destroyed.

**3.53 data structure:** A set of data items and the relationship between them representing the whole or a part of a constituent.

NOTE – The data items constituting a data structure represent attributes of the document, the document profile, the components, the styles, or the content portions concerned.

**3.54 description:** A constituent that corresponds to a structural element.

NOTE – A basic object may have several descriptions when alternative descriptions are used.

**3.55 descriptor:** A data structure representing the document profile, an object class description, a layout style, a presentation style, an object description, or a protected part description.

**3.56 device space:** A colour space where the coordinates are those used by a particular device in the measurement or rendition of colour.

**3.57 digital signature:** A form of seal associated with a specified part of a document which provides proof of uniqueness of the identity of the originator (source) who applied the seal; it supports non-repudiation of origin of the sealed (signed) part.

**3.58 document:** A structured amount of information intended for human perception, that may be interchanged as a unit between users and/or systems.

**3.59 document application profile:** The specification of a combination of features defined in ITU-T Rec. T.410-Series | ISO/IEC 8613, intended to form a subset to fulfil the requirements of an application.

**3.60 document architecture:**

- a) Rules for defining the structure of documents, in terms of collections of constituents and the attributes of which these are composed.
- b) The structural information of a document consisting of the set of one or more of the following structures: specific logical structure, specific layout structure, generic logical structure and/or generic layout structure together with presentation styles and/or layout styles.

**3.61 document architecture class:** The rules for defining the structure and representation of documents in formatted form, processable form or formatted processable form.

**3.62 document body:** The part of a document that may include a generic logical and layout structure, specific logical and layout structure, layout and presentation styles, protected parts but excludes the document profile.

**3.63 document class:** A set of logical object class descriptions, layout object class descriptions, generic content portion descriptions, styles and a document profile, that specifies a set of documents with common characteristics.

**3.64 document class description:** The specification of a document class.

**3.65 document layout process:** The process that creates a specific layout structure in accordance with the generic layout structure and information contained in the specific logical structure, the generic logical structure and the layout styles.

**3.66 document layout root:** The composite object of the specific layout structure at the highest level of the hierarchy.

- 3.67 document logical root:** The composite object of the specific logical structure at the highest level of the hierarchy.
- 3.68 document profile:** A set of attributes which specifies the characteristics of the document as a whole.
- 3.69 editing:** The carrying out of operations associated with creation and modification of the structure and/or the content of a document.
- 3.70 editing process:** The stage of document processing that consists of the content editing process and the logical structure editing process.
- 3.71 end-aligned:**
- a) The result of a layout or imaging process that positions the sequence of character images for a line such that the escapement point of the last character image is positioned at the end edge of the positioning area.
  - b) A tabulation alignment that positions the sequence of character images for a specified character string such that the escapement point of the last character image is positioned at the tabulation stop.
- 3.72 end edge:** The edge of the positioning area of a basic layout object that is in the direction of the character path.
- 3.73 escapement point:** A reference point associated with a character image that is used for positioning of the next character image.
- 3.74 external document class:** A document class referred to by the document profile of an interchanged document containing no generic structure.
- 3.75 factor set:** One or more object class descriptions which are used to factorise the attributes of object descriptions representing a specific structure.
- 3.76 file:** A named collection of information.
- 3.77 file section:** For a file recorded over more than one volume, that part of the file that is recorded on any one volume.
- NOTE – The term *file* in ISO 9293 is equivalent to the term *file section* in this Specification.
- 3.78 file space:** The region within which a file section may be recorded on a volume.
- 3.79 filing:** The storage of a document according to some defined method in order to facilitate retrieval of the document.
- 3.80 fingerprint:** A short and compact code that may be computed in order to characterize some specified information, with the property that it is not practicable to construct different information which would yield the same output.
- 3.81 font:** A set of character images normally with a common design and size.
- 3.82 font size:** The height of the character images in a font.
- 3.83 formatted form:** A form of representation of a document that allows the presentation of the document as intended by the originator and that does not support editing and (re)formatting.
- 3.84 formatted processable form:** A form of representation of the document that allows presentation of the document as intended by the originator and also supports editing and (re)formatting.
- 3.85 formatting:** The carrying out of operations to determine the layout of a document.
- 3.86 frame:** A type of composite layout component that corresponds to a rectangular area within a page or another frame.
- 3.87 generic content portion:** A content portion associated with an object class.
- 3.88 generic content portion description:** A content portion description associated with an object class description.
- 3.89 generic-document:** A structured amount of information intended for the interchange of generic structures, and optionally associated styles and content portions, for use in the processing of interchanged documents.
- 3.90 generic layout structure:** A set of layout object classes and associated generic content portions.
- 3.91 generic logical structure:** A set of logical object classes and associated generic content portions.

- 3.92 geometric graphic element:** A graphic element used to describe an image by geometric graphical means.  
NOTE – Geometric graphics elements include those describing primitive geometric shapes such as points, arcs, lines.
- 3.93 graphic character:** A member of a set of graphic symbols used for the representation of information.  
NOTE – Graphic characters include simple alphanumeric characters (for example, accented letters) and pictorial characters (for example, mosaics).
- 3.94 graphic element:** A content element that is capable of having a visual representation.  
NOTE – Three types of graphic elements are distinguished in the ITU-T Rec. T.410-Series | ISO/IEC 8613: graphic characters, raster graphics elements, and geometric graphics elements.
- 3.95 hard line terminator:** A line terminator that is intended not to be removed in a reformatting process.
- 3.96 horizontal direction (of a layout object):** The direction in a layout object relative to which content architectures may define attributes determined by using the horizontal axis of the page.
- 3.97 imaging order:** The order of precedence of layout objects for imaging in the layout object to which they are immediately subordinate.
- 3.98 imaging process:** The process of producing a document on a presentation medium in human-perceptible form, making use of the document profile, specific and generic layout structures, presentation styles and content portions.
- 3.99 implicit layout category:** The layout category of a logical object which is used in the absence of an explicit specification of a layout category for that object.
- 3.100 indentation:** The result of a layout or imaging process that causes the sequence of character images for a line to begin at a distance from the line home position in the direction of the character path.
- 3.101 initial point:**
- a) The point associated with a basic layout object relative to which all line boxes imaged within that basic layout object are positioned (character content architectures ITU-T Rec. T.416 | ISO/IEC 8613-6).
  - b) The point associated with a basic layout object relative to which all pels imaged within that basic layout object are positioned (raster graphics content architectures ITU-T Rec. T.417 | ISO/IEC 8613-7).
- 3.102 integrity:** Used here synonymously with data integrity.
- 3.103 intended recipient:** A recipient of a document that is expected to receive or have access to the document.
- 3.104 interchange:** The process of providing a document to a receiving person or device, by means of data communication or by exchange of storage media.
- 3.105 interchange data element:** A data structure representing a constituent of a document.
- 3.106 interchange format:** The rules for representing a document for the purpose of interchange.
- 3.107 interchange format class:** A form of interchange format suitable to a specific application.  
NOTE – In the ITU-T Rec. T.410-Series | ISO/IEC 8613, the defined classes differ by the ordering of the interchange data elements or by the coding.
- 3.108 interchangeable storage medium:** Storage medium which can be used to interchange information by moving the medium from one information processing system to another.
- 3.109 inter-character space:** An additional amount of spacing that is included between adjacent character images.
- 3.110 intersection:** The common area of two or more layout objects that overlap each other partially or fully on the presentation medium.
- 3.111 item identifier:** A string of characters that precedes and is separated from the remainder of the first line of a basic component with content. An item identifier is used to identify the subsequent text.
- 3.112 justified:** The result of a layout or imaging process that varies the width of the space character and/or the inter-character space to produce a simultaneously start-aligned and end-aligned presentation of the text.
- 3.113 kern:** The part of a character image which extends beyond its position point or escapement point.
- 3.114 layout category:** The association of basic logical objects with lowest level frames such that the content of these basic logical objects is placed in the appropriate frames.
- 3.115 layout object:** An element of the specific layout structure of a document, for example, page, block.

**3.116 layout object class:** An element of the generic layout structure from which a set of layout objects with common characteristics may be derived, for example, pages with common headers and footers.

**3.117 layout process:** The stage of document processing that consists of the document layout process and the content layout process.

NOTE – This is also referred to as formatting.

**3.118 layout stream:** A set of basic logical objects pertaining to the same layout category.

**3.119 layout structure:**

- a) The result of dividing and subdividing the content of a document into increasingly smaller parts, on the basis of the presentation, for example, into pages, blocks.
- b) All layout objects and associated content portions forming the layout hierarchy of a document.

**3.120 layout style:** A constituent of the document, referred to from a logical component, that guides the creation of a specific layout structure.

**3.121 leading edge:** The edge of a frame or block that is orthogonal to the direction of the layout path and that is met first, from the outside of the frame or the block, in the opposite direction of the layout path.

**3.122 left hand edge:** The edge of a frame or block that is parallel to the direction of the layout path and that is met first, from the outside of the frame or the block, in the direction at an angle of 270 degrees counterclockwise relative to the direction of the layout path.

**3.123 line box:** A rectangular area within which a sequence of character images are positioned.

**3.124 line home position:** The point within a line box that is used for positioning that line box.

**3.125 line progression:**

- a) The direction of progression of successive line boxes within a basic layout object (character content architectures ITU-T Rec. T.416 | ISO/IEC 8613-6).
- b) The direction of progression of successive lines of pels within a basic layout object (raster graphics content architectures ITU-T Rec. T.417 | ISO/IEC 8613-7).

**3.126 line spacing:**

- a) The distance between two adjacent reference lines within a basic layout object (character content architectures ITU-T Rec. T.416 | ISO/IEC 8613-6).
- b) The distance between two adjacent lines of pels within a basic layout object (raster graphics content architectures ITU-T Rec. T.417 | ISO/IEC 8613-7).

**3.127 line terminator:** A control function or combination of control functions that indicates the end of a line or the end of a character sequence.

**3.128 logical object:** An element of the specific logical structure of a document which may have a meaning that is significant to the application or user, for example, chapter, section, paragraph.

**3.129 logical object class:** An element of the generic logical structure from which a set of logical objects with common characteristics may be derived, for example, composite logical objects representing sections with a common internal structure.

**3.130 logical structure:**

- a) The result of dividing and subdividing the content of a document into increasingly smaller parts, on the basis of the human-perceptible meaning of the content, for example, into chapters, sections, paragraphs.
- b) All logical objects and associated content portions representing the logical hierarchy of a document.

**3.131 logical structure editing process:** The process that creates a new specific logical structure or modifies a previous specific logical structure and allocates or (re)allocates content to basic logical objects.

**3.132 mandatory attribute:** An attribute which, when applicable to a constituent, must be specified explicitly in the constituent.

**3.133 nominal page:** A rectangular area which, as assumed by the sender of a document, has the ideal size of the presentation surface.

NOTE – Examples of ideal sizes are given in ISO 216.

**3.134 non-basic:** A qualifier for attribute values, control function parameter values and other capabilities that are allowed in document interchange in the context of a given document application profile only if their use is declared in the document profile.

**3.135 non-mandatory attribute:** An attribute which, when applicable to a constituent, need not be specified explicitly; if the attribute is not specified explicitly in a given constituent, the attribute does not apply.

**3.136 non-repudiation of origin:** The property that an originator can be proved to the satisfaction of a third party to be the source of a document or a specified part of it.

**3.137 object:** An element of the specific layout structure or of the specific logical structure.

**3.138 object class:** An element of a generic structure from which objects with common characteristics may be derived.

**3.139 object class description:** A set of attributes that specify the properties of an object class including its relationships, if any, with other components.

**3.140 object description:** A set of attributes that specify the properties of an object including its relationships, if any, with other components.

**3.141 object type:** A property of every component that specifies which attributes are permitted in the description to which it applies and indicates the role of the component in the document architecture.

**3.142 ODA data stream:** A data stream conforming to the ITU-T Rec. T.410-Series | ISO/IEC 8613 in which the data elements representing constituents and attribute values are in accordance with ITU-T Rec. T.412, T.414 ISO/IEC 8613-2, 8613-4 and one or more of the ITU-T Rec. T.410-Series | ISO/IEC 8613 Specifications including:

- ITU-T Rec. T.416 | ISO/IEC 8613-6;
- ITU-T Rec. T.417 | ISO/IEC 8613-7;
- ITU-T Rec. T.418 | ISO/IEC 8613-8;

and any referenced standards.

**3.143 Open Document Language; ODL (abbreviation):** A Standard Generalized Markup Language (SGML, ISO 8879) application for representing documents conforming to ISO/IEC 8613.

**3.144 orphan:** One or more lines of text that are associated with subsequent text but isolated from it by a page or column boundary.

**3.145 overhang:** The result of a layout or imaging process that positions the sequence of character images for a line to begin at a distance from the line home position in the direction opposite to the character path.

**3.146 page:** A layout component that corresponds to a rectangular area used for presenting the content of the document.

**3.147 page coordinate system:** An orthogonal coordinate system whose origin is the top left corner of the page; its horizontal axis and its vertical axis coincide with the top edge and the left edge of the page, respectively.

**3.148 page set:** A composite layout component that represents a collection of pages or further page sets.

**3.149 pairwise kerning:** The distance between two adjacent character images depending on the combination of the two characters together rather than separately.

**3.150 parallel annotation:** Two sequential character strings that are presented in parallel, wherein the second string is used to indicate the pronunciation or interpretation of the first string.

**3.151 partial generator set:** A collection consisting of hierarchically related object class descriptions which are used to guide the creation of hierarchically related corresponding object descriptions but do not fully specify all specific structures that may be created.

**3.152 pel array:** A two-dimensional array of pels used to represent a pictorial image.

**3.153 pel path:** The direction of progression of successive pels along a line within a basic layout object.

**3.154 pel spacing:** The distance between any two successive pels along a line within a basic layout object.

**3.155 picture element; pel (abbreviation):** The smallest graphic element that may be individually addressed within a picture (an alternative term for raster graphics element).

- 3.156 positioning area:** The rectangular area within a basic layout object within which the position points and the escapement points of all character images are located.
- 3.157 position point:** The point relative to which the character image is placed, i.e. the character is imaged with the position point at the active position.
- 3.158 presentation:** The operation of rendering the content of a document in a form perceptible to a human being.
- 3.159 presentation medium:** The carrier of information in a form perceptible to a human being.
- 3.160 presentation style:** A constituent of the document, referred to from a basic logical or layout component, which guides the format and appearance of the document content.
- 3.161 presentation surface:** A two-dimensional presentation medium (such as paper, film, video display screen) on which the formatted form of a document may be displayed for human viewing.
- 3.162 primary description:** A description that represents a basic object that best represents the intent of the originator.
- 3.163 primary subtree:** The basic object description in conjunction with its associated content portion descriptions, with the potential to be replaced by an alternative subtree.
- 3.164 privileged recipient:** A recipient of a document that in addition to being an intended recipient, has the right to perform certain security-related operations intended for that particular recipient, such as to interpret specified enciphered parts of the document, and to perform integrity and authenticity checks on specified parts of the document.
- 3.165 processable form:** A form of representation of a document that allows editing and formatting.
- 3.166 processing:** The carrying out of operations on a document, including editing, (re)formatting, presentation, filing and retrieval.
- 3.167 protected part:** A constituent consisting of a part of the document that is sealed or enciphered, e.g. a sealed document profile or pre-enciphered document body part.
- 3.168 raster graphics element:** An alternative term for a picture element (pel).
- 3.169 recipient:** Any application or user receiving or having access to a document.
- 3.170 record:** A sequence of bytes treated as a unit of information stored on a file.
- standard for volume and file structure: An international standard that specifies the format of the recorded structures that contain descriptive information about the volume and the files recorded on an interchangeable storage medium.
- 3.171 reference area:** A rectangular area within a basic layout object, with its sides equal to the pel spacing and the line spacing, within which the main part of a pel is imaged.
- 3.172 reference colour space:** Basic colour space within ODA relative to which relationships to colour spaces used in document and content architectures are specified.
- NOTE – The reference colour space is defined by the Commission Internationale de l'Eclairage (CIE) Standard Colorimetric System (XYZ), 1931.
- 3.173 reference line:** A line through the line home position and parallel to the character path.
- 3.174 reference point:** The point at the corner of the reference area situated in the opposite direction of both pel path and line progression and which is used for positioning a pel.
- 3.175 reformatting:** The carrying out of operations to determine the new layout of a previously formatted document.
- 3.176 region of interest:** A rectangular area within a virtual coordinate space, with sides which are parallel to the axes of its coordinate system, surrounding those (parts of) geometric graphics that are intended to be imaged.
- 3.177 resource-document:** A generic-document containing one or more object class descriptions and their associated generic content descriptions referred to by one or more object class descriptions of another document.
- 3.178 retrieval:** The recovery of previously filed information.
- 3.179 Red Green Blue; RGB (abbreviation) colour space:** A colour space based on red, green and blue stimuli or primaries.

#### NOTES

- 1 The RGB values are not to be confused with the CIE  $\overline{rgb}$  as defined in CIE colorimetry.
- 2 RGB values are negative in certain areas outside the gamut defined by the RGB primaries.

**3.180 right hand edge:** The edge of a frame or block that is parallel to the direction of the layout path and that is met first, from the outside of the frame or the block, in the direction at an angle of 90 degrees counterclockwise relative to the direction of the layout path.

**3.181 scaled measurement unit; SMU (abbreviation):** A unit of linear measurement used for positioning and dimensioning layout objects and content elements on a presentation surface, its value being equal to the basic measurement unit (BMU) times the unit scaling.

**3.182 seal (noun):** Data associated with a specified part of a document by an originator, which a recipient (privileged recipient) may use to verify the integrity and authenticity of the specified part.

**3.183 seal (verb):** To associate a seal with a specified part of a document.

**3.184 security domain:** The set of resources subject to a single security policy.

**3.185 security label:** A marking of a document, which specifies the handling of the document according to the security policy in force.

**3.186 security policy:** The set of rules that specify the procedures and services required to maintain the intended level of security of a set of resources.

**3.187 sequential layout order:** The sequential order in which layout objects are to be imaged when an imaging order is not specified.

**3.188 sequential logical order:** The sequential order in which logical objects are to be processed by the layout process.

**3.189 sequential order:** A convention for ordering the objects in a structure such that each object is succeeded by all of its immediate subordinates, before any other object with the same immediate superior, i.e. in order tree traversal.

**3.190 soft line terminator:** A line terminator that is allowed to be removed, relocated or replaced in a subsequent layout process.

**3.191 spacing ratio:** The ratio of line spacing to pel spacing.

**3.192 specific layout structure:** A set of layout objects and associated content portions.

**3.193 specific logical structure:** A set of logical objects and associated content portions.

**3.194 start-aligned:**

- a) The result of a layout or imaging process that positions the sequence of character images for a line such that the position point of the first character image of that sequence is positioned on the line home position or at the point specified by the first line indentation or overhang, if any, for the line.
- b) A tabulation alignment that positions the sequence of character images for a specified character string such that the position point of the first character image of that sequence is positioned at the tabulation stop.

**3.195 start edge:** The edge of the positioning area of a basic layout object that is in the direction opposite to the character path.

**3.196 stream root category:** A property of components that can be used to designate the layout objects into which the content associated with logical objects shall be placed.

**3.197 stream root sub category:** A property of components that can be used to designate the subordinate layout objects into which the content associated with logical objects shall be placed. The designated subordinate layout objects shall all be subordinates of a single layout object that has the property of being a stream root category layout object.

**3.198 structural element:** The structural elements of a document are the content portion, the object and the object class.

**3.199 tabulation alignment:** A layout or imaging process that causes the sequence of character images for a specified character string to be positioned according to a specified method, (start-aligned, end-aligned, centred or aligned-around) at a specified point (tabulation stop) along a reference line.

**3.200 tabulation stop:** A position along a reference line that is to be used for a specified method (start-aligned, end-aligned, centred or aligned-around) of tabulation alignment.

**3.201 tile:** One element of a two dimensional array of non-overlapping rectangular regions of a pel array.

**3.202 text unit:** A data structure representing a content portion description.

- 3.203 top edge:** The edge of the positioning area of a basic layout object that is in the direction opposite to the line progression.
- 3.204 top left corner:** The corner of a layout object that is least progressed both in the horizontal and vertical directions of this layout object.
- 3.205 top right corner:** The corner of a layout object that is most progressed in the horizontal direction and least progressed in the vertical direction of this layout object.
- 3.206 trailing edge:** The edge of a frame or block that is orthogonal to the direction of the layout path and that is met first, from the outside of the frame or the block, in the direction of the layout path.
- 3.207 unit scaling:** A scaling factor (an integer or a fraction) that is applied to the basic measurement unit (BMU) to derive a scaled measurement unit (SMU).
- 3.208 variable spacing:** The characteristics of a font wherein the distance between the position point and the escapement point for different character images may be different.
- 3.209 volume:** A dismountable physical unit of storage medium, for example a flexible disk cartridge.
- 3.210 widow:** One or more lines of text associated with preceding text but isolated from it by a page or column boundary.

## 4 Abbreviations

For the purposes of this Recommendation | International Standard, the following definitions apply.

|        |                                                                                                 |
|--------|-------------------------------------------------------------------------------------------------|
| ASN.1  | Abstract Syntax Notation One                                                                    |
| BMU    | Basic Measurement Unit                                                                          |
| BNF    | Backus-Naur-Form                                                                                |
| CF     | Character Formatted                                                                             |
| CGM    | Computer Graphics Metafile                                                                      |
| CIE    | Commission Internationale de l'Eclairage                                                        |
| CIELAB | Commission Internationale de l'Eclairage lightness-hue-chroma CIE 1976 $L^*a^*b^*$ colour space |
| CIELUV | Commission Internationale de l'Eclairage lightness-hue-chroma CIE 1976 $L^*u^*v^*$ colour space |
| CMY    | Cyan Magenta Yellow                                                                             |
| CMY(K) | Cyan Magenta Yellow Black                                                                       |
| DAPPN  | Document Application Profile Proforma Notation                                                  |
| DIS    | Disallowed                                                                                      |
| ECMA   | European Computer Manufacturer's Association                                                    |
| FDA    | Formatted Document Architecture                                                                 |
| FODA   | Formal Specification of the Open Document Architecture                                          |
| FPDA   | Formatted Processable Document Architecture                                                     |
| ID     | Identifier                                                                                      |
| IMCL   | Information Modeling by Composition Language                                                    |
| IPM    | Interpersonal Message                                                                           |
| IPMS   | Interpersonal Message System                                                                    |
| LUT    | Colour LookUp Table                                                                             |
| MHS    | Message Handling System                                                                         |
| mm     | Millimeter                                                                                      |
| MOTIS  | Message Oriented Text Interchange System                                                        |
| MTS    | Message Transfer System                                                                         |
| MUL    | Multiple (Repetition)                                                                           |
| NIST   | National Institute of Standards and Technology                                                  |
| NOVR   | No Override                                                                                     |

|       |                                                    |
|-------|----------------------------------------------------|
| ODA   | Open Document Architecture                         |
| ODIF  | Open Document Interchange Format                   |
| ODL   | Open Document Language                             |
| PDA   | Processable Document Architecture                  |
| pel   | Picture Element                                    |
| PERM  | Permitted                                          |
| PMUL  | Permitted Multiple (Optional Repetition)           |
| RGB   | Red Green Blue                                     |
| REQ   | Required                                           |
| SDIF  | SGML Document Interchange Format                   |
| SGML  | Standard Generalized Markup Language               |
| SMPTE | Society of Motion Picture and Television Engineers |
| SMU   | Scaled Measurement Unit                            |
| tpi   | Tracks Per Inch                                    |
| tpmm  | Tracks Per Millimeter                              |

## 5 Conventions

The following conventions are used within this Recommendation | International Standard. In this clause, the word parameter(s) is to be interpreted to apply to parameter(s), sub-parameter(s), sub-sub-parameter(s), etc. to any level.

### 5.1 Names of attributes and parameters

The names of attributes and parameters are referenced by giving the name of the attribute or parameter in double quotation marks preceded by the word attribute or parameter. For example:

- ...expressed by the attribute “subordinates”. This attribute...;
- ...two parameters “horizontal dimension”...;
- ...by a sub-parameter “fixed dimension”....

In some cases terms are used to describe concepts which have the same name as an attribute or parameter. In the case of reference to such concepts neither the word attribute or parameter nor quotes are used.

### 5.2 Names of attribute and parameter values

If the values of attributes and parameters are terminal symbols, they are referenced by giving the name of the attribute or parameter value in single quotation marks. For example:

- ...value is ‘opaque’;
- ...this parameter allows a sender to select a recto or a verso presentation of the page by specifying the value ‘recto’ or ‘verso’, respectively.

### 5.3 Component and component description

The term component may be used in conjunction with qualifying terms, including: basic, composite, page set, page, frame, block, logical and layout. For example:

- a) layout components containing blocks means layout objects containing blocks, or layout object classes containing blocks;
- b) frame component means an object of type frame, or a class to which objects of type frame belong.

The term component description may be qualified in the same way.

### 5.4 Introduction of terms

Italics are used for the purpose of highlighting the point at which the definition of a term occurs in the text.

These terms are also defined in clause 3.

## 5.5 Representation of degrees

A lower case d is used for the purpose of signifying that the associated value specifies degrees. For example, the notation d0 specifies zero degrees and d180 specifies 180 degrees.

# 6 General concept of ODA

## 6.1 Purpose of ODA

The purpose of the document architecture is to facilitate the interchange of documents in a manner such that:

- different types of content, including text, image, graphic and audio, can coexist within a document;
- the intentions of a document originator with respect to the logical and layout structure of a document can be transmitted to a recipient.

It also supports security aspects such as

- protection against unauthorized semantic knowledge of parts of a document;
- detection of discrepancies between the claimed and actual source and content of parts of the document.

This clause uses a number of terms for which definitions are given in clause 3. However, for the purpose of the current clause, different though compatible definitions of the essential terms are given below.

- *Interchange* – The process of providing a document to a receiving person or device, by means of data communication or by exchange of storage media.
- *Editing* – The carrying out of operations associated with creation and modification of the structure and/or the content of a document.
- *Formatting* – The carrying out of operations to determine the layout of a document, i.e. the appearance of its content on a presentation medium.
- *Presentation* – The operation of rendering the content of a document in a form perceptible to a human being. Typical presentation media are paper, and video screens.

The document architecture provides for the representation of documents in three forms:

- formatted form, that allows documents to be presented as intended by the originator;
- processable form, that allows documents to be edited and formatted;
- formatted processable form, that allows documents to be presented as well as edited and reformatted.

Alternative terms commonly used are final form and image form for formatted form, and revisable form for processable form.

Each of these forms allows the originator to express intentions regarding the structuring and/or formatting of the interchanged document.

## 6.2 Overall concept of ODA

The concept of ODA is based on

- the existence of a layout view and a logical view of the document; the view from the physical viewpoint (for example, a collection of pages) and the view in the sense of its abstract components (for example, an assembly of chapters);
- the existence of a specific structure and a generic structure; the specific document structure is the one that the user may read; the generic structure is the template that guides the creation of the document and that could be re-used for its amendment;
- the existence of document classes; a document class is the set of generic features that are common to a category of documents (for example, Sales Report Form).

### 6.2.1 Logical structure and layout structure

The key concept in the document architecture is that of structure. Document structure is the division and repeated subdivision of the content of a document into increasingly smaller parts. The parts are called objects. The structure has the form of a tree.

The document architecture permits two structures to be applied to a document: a logical structure and a layout structure. Any one or both structures may be applied to a given document.

In the logical structure, the document is divided and subdivided on the basis of the meaning. Examples of logical objects are chapters, sections, figures and paragraphs.

In the layout structure, the document is divided and subdivided on the basis of the layout. Examples of layout objects are pages and blocks.

An example of the logical view of a document called report is shown in Figure 1.

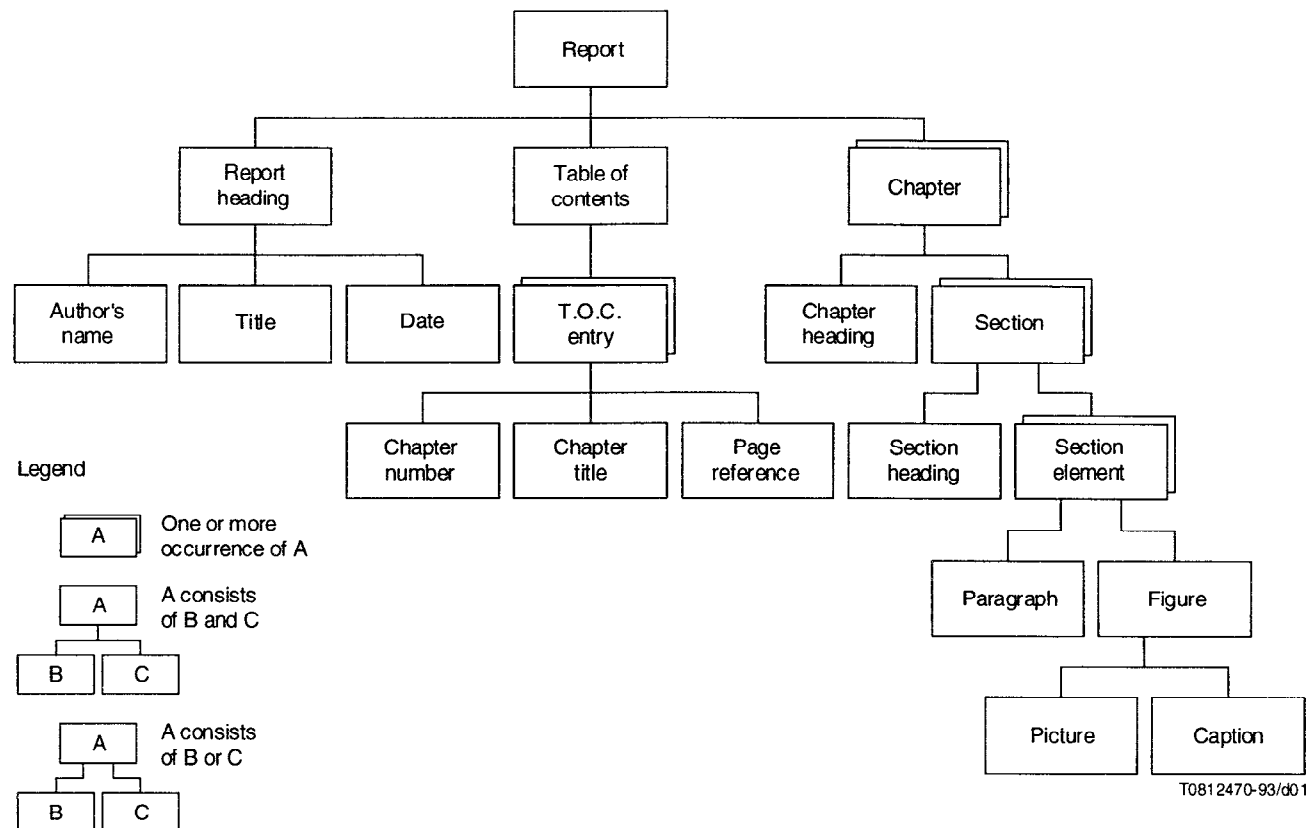


**Figure 1 – Example of a logical view of a document**

The logical structure and the layout structure provide alternative but complementary views of the same document. For example, a book can be regarded as consisting of chapters containing figures and paragraphs, or alternatively, as consisting of pages that contain text blocks and/or graphic blocks.

An object that is not subdivided into smaller objects is called a basic object. All other objects are called composite objects.

The following types of layout objects are defined in the document architecture:

- *block* – A basic layout object corresponding to a rectangular area on the presentation medium containing a portion of the document content.
- *frame* – A composite layout object corresponding to a rectangular area on the presentation medium and containing either one or more frames or one or more blocks.
- *page* – A basic or composite layout object corresponding to a rectangular area on the presentation medium and, if it is a composite object, containing either one or more frames or one or more blocks or, if it is a basic object, containing one or more content portions of the document content.
- *page set* – A set of one or more page sets and/or pages.
- *document layout root* – The highest level object in the hierarchy of the specific layout structure.

For logical objects, no classification other than basic logical object, composite logical object and document logical root is defined in the document architecture. Logical object categories such as chapter, section and paragraph are application-dependent and can be defined by a document application profile using the object class mechanism (see 6.2.6).

### 6.2.2 Content portions

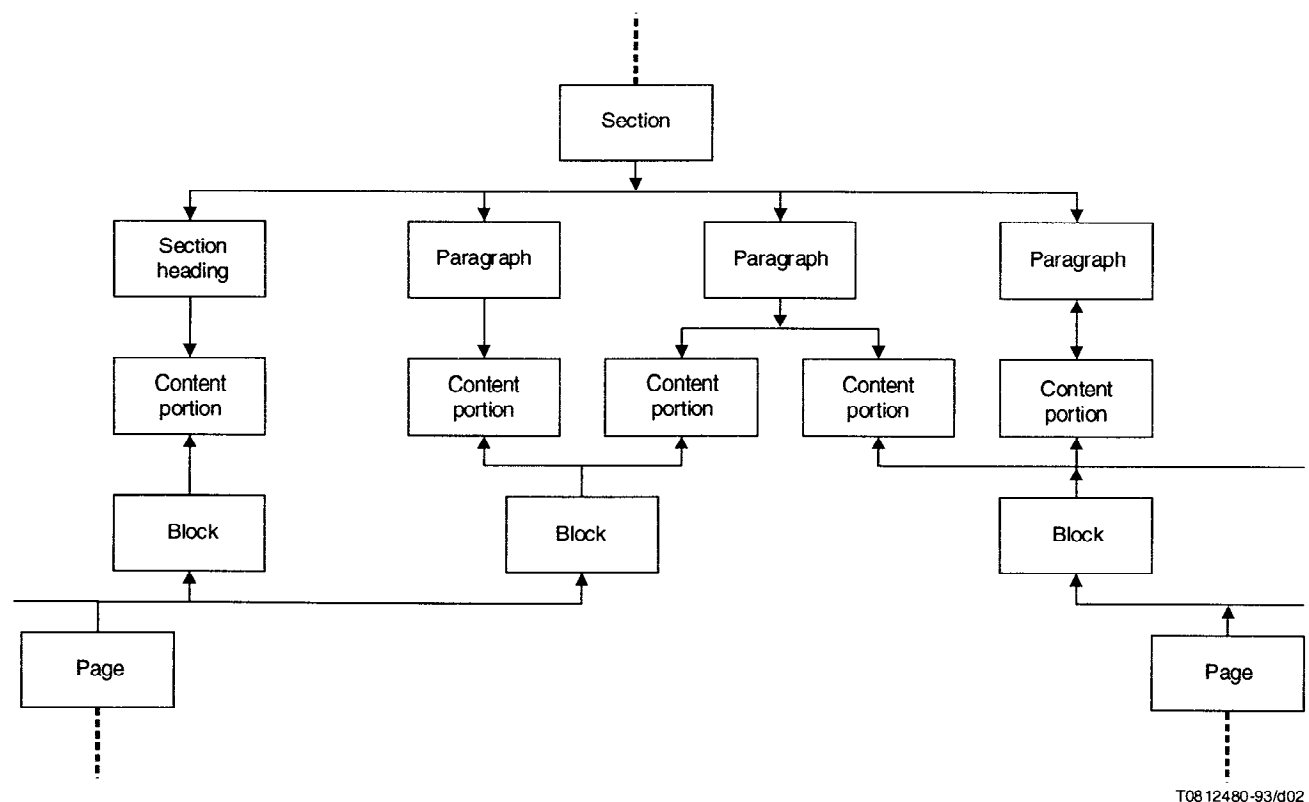
The basic elements of the content of a document are called content elements. For content consisting of character text, the content elements are characters. In the case of images or graphics, the content elements are picture elements (also called pels) or geometric graphics elements (lines, arcs, polygons, etc.).

When a document has both logical structure and layout structure, each content element belongs, in general, to exactly one basic logical object and to exactly one basic layout object. A set of related content elements that belong to one basic logical object (if the document has any logical structure) and one basic layout object (if the document has any layout structure) is collected within a constituent type called a content portion.

The relationships between basic logical objects, basic layout objects and content portions are:

- a basic logical object has associated with it one or more content portions;
- a basic layout object has associated with it one or more content portions;
- any logical or layout object (basic or composite) has associated with it zero or more content portions;
- there is, in general, no one-to-one correspondence between logical objects and layout objects.

The last point is illustrated by Figure 2.



T08 12480-93/d02

Figure 2 – Possible correspondence between logical and layout objects

### 6.2.3 Content architectures

A content portion associated with a basic logical object or a basic layout object may have a more detailed internal structure. The rules governing such an internal structure depend on the type of content and are called a content architecture. The content of a basic logical object or a basic layout object is structured according to only one content architecture.

Definitions of content architectures are given in other Specifications of the ITU-T Rec. T.410-Series | ISO/IEC 8613. These include:

- ITU-T Rec. T.416 | ISO/IEC 8613-6: Character content architectures;
- ITU-T Rec. T.417 | ISO/IEC 8613-7: Raster graphics content architectures;
- ITU-T Rec. T.418 | ISO/IEC 8613-8: Geometric graphics content architectures.

#### 6.2.4 Attributes

An attribute is a property of a document, or of a document constituent (i.e. a logical object, a layout object, a logical object class, a layout object class, a style, a content portion or a protected part). It expresses a characteristic of the document or document component concerned, or a relationship with one or more documents or document components.

The set of attributes associated with a document as a whole is called a document profile. It represents reference information about the document and may repeat information in the document content, for example, the title and the name of the author.

The set of attributes that applies to a logical object or a layout object depends on the type of the object; different sets of attributes are defined for basic logical objects, composite logical objects, document logical root, blocks, frames, pages, page sets and document layout root. These are called document architecture attributes. In general, document architecture attributes are independent of the type of content of the objects to which they apply.

Examples of document architecture attributes are:

- the attribute “object identifier” (all objects);
- the attribute “subordinates” (composite objects);
- layout directives such as the attribute “concatenation”, the attribute “offset”, the attribute “separation” (basic logical objects);
- the attribute “position” (blocks and frames);
- the attribute “dimensions” (blocks, frames and pages).

In addition to the document architecture attributes, a set of presentation attributes applies to basic logical and basic layout objects. The set of presentation attributes that applies to a given basic object depends on the content architecture governing the content of this object; a different set of presentation attributes is defined for each content architecture.

Examples of presentation attributes are:

- the attribute “character spacing” (character content architectures);
- the attribute “clipping” (raster graphics content architectures);
- the attribute “line rendition” (geometric graphics content architectures).

Presentation attributes may be collected into presentation styles, to which references may be made from both logical and layout objects. The attributes that apply to a content portion include a content portion identifier and a set of coding attributes, the composition of which depends on the coding method used for the content, for example, the attribute “number of pels per line” for facsimile-coded raster graphics images.

#### 6.2.5 Relations between logical structure and layout structure

The logical structure and the layout structure are, in principle, independent of each other. The logical structure of a document is determined by the author and embedded in the document during the editing process. The layout structure is usually determined by a formatting process. In addition to the presentation attributes, the formatting process may be controlled by attributes called layout directives associated with the logical structure.

Examples of layout directives are:

- the requirement that a chapter starts on a new page;
- the requirement that the title of a section and the first two lines of its first paragraph are presented on the same page.

Layout directives may be collected into layout styles each of which may be referred to by one or more logical objects.

#### 6.2.6 Specific and generic structures

In a document, the logical objects and/or the layout objects can often be classified into groups of similar objects. Therefore the concept of object class is introduced.

The similarity can be related to logical features such as chapter, section or paragraph hierarchy, to layout features such as size or style, or to content such as page headers and footers. Even an entire document may be a member of a group of similar documents, a letter, a memorandum or a report.

An object class or a document class is a specification of the set of properties that are common to its members. Such a specification consists of a set of rules to determine the values of the attributes that specify the common properties. These rules are used to control the consistency among the objects or documents making up the class, and to facilitate the creation of additional objects or documents within the class.

The set of logical object classes and layout object classes associated with a document, and their relationships, are called generic logical structure and generic layout structure.

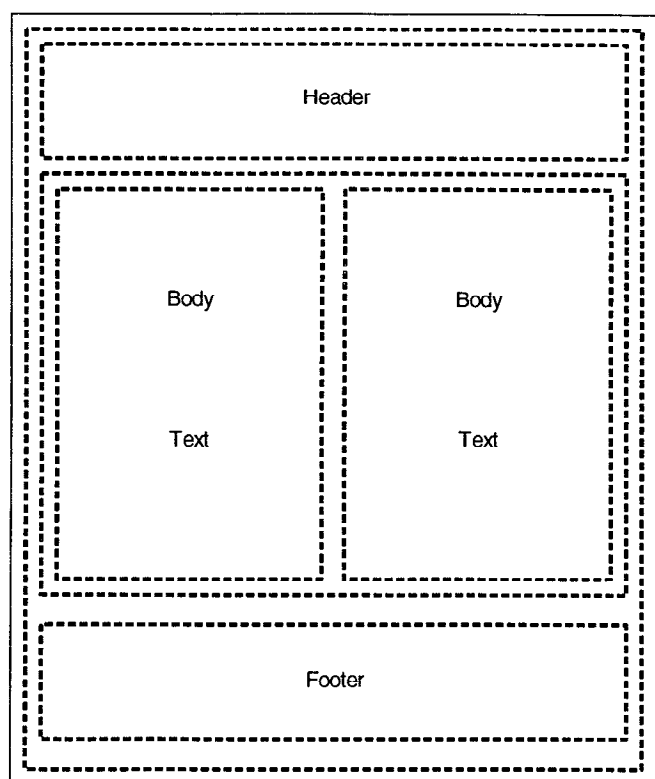
The structures that are particular to a given document are named specific logical structure and specific layout structure.

A document class contains a generic logical structure and/or a generic layout structure. In addition, all document class descriptions shall include a document profile and may include layout styles, presentation styles and generic content portion descriptions.

The generic logical structure represents the set of all potential specific logical structures, and the generic layout structure represents the set of all potential specific layout structures that are applicable to the document class concerned.

The generic logical structure is used as a set of rules from which specific logical objects and structures are derived during the editing process. The generic layout structure is used as a set of rules from which specific layout objects and structures are derived during the formatting process.

An example of generic layout structure is depicted in Figure 3 which shows a page layout with frames for a header, a footer and two columns of body text.



T0812490-93/d03

Figure 3 – Example of a page layout

### 6.2.7 Document profile

The document profile consists of a set of attributes associated with a document as a whole. In addition to reference information such as title, date and author's name, which facilitates storage and retrieval of the document, the document profile contains a summary of the document architecture features some of which are and the remainder of which may be used in the document, in order that a recipient can easily determine which capabilities are required for processing or imaging the document. The attributes representing the latter type of information are called document characteristics and include:

- a specification of the form (formatted, processable or formatted processable) of the document;
- specifications of the content architectures used in the document;
- specifications of the character sets, character fonts, character styles, character orientations and types of emphasis used in the document.

The document profile may be interchanged alone.

### 6.2.8 Generic-document

A generic-document consisting of a document profile and generic structures may be used to assist in the processing of interchanged documents. A generic-document may be interchanged.

### 6.2.9 Protected parts

Protected parts enable protection to be provided for parts of a document offering confidentiality, integrity, authenticity and non-repudiation of origin.

Protected parts cannot provide for the protection of a document as a whole. Treatment of the complete document as an object is outside the scope of this series of Recommendations | International Standards.

Parts of a document may be protected before the layout process in their processable form, or after the layout process in formatted or formatted processable forms.

## 6.3 Document processing model

### 6.3.1 Relationships of ODA to document processing

The ITU-T Rec. T.410-Series | ISO/IEC 8613 is concerned with the definition of a document architecture which permits processing of interchanged documents. A model of document processing is provided as a basis for determining the scope of the processes described by the ITU-T Rec. T.410-Series | ISO/IEC 8613.

A basic model of document processing is summarized in this subclause (see Figure 4). This model is not intended to represent an actual implementation, nor to restrict in any way the processing that may be applied to an interchanged document.

Conceptually, a document is viewed as progressing through three phases of processing as shown in Figure 4. The order of the processes is not intended to imply that they are performed sequentially in an actual implementation.

### 6.3.2 Editing process

The document editing process is concerned with creating a new document or modifying an existing one. The document architecture provides data structures for representing the document resulting from this process.

While document creation and modification can differ in the functions performed and procedures followed, they are considered to be equivalent in the view of this model because the result of both is the same: a new document.

Upon completion of editing, the resulting document may be interchanged. Such a document is said to be interchanged in processable form; it is suitable for input to either the editing or layout process.

### 6.3.3 Layout process

The document layout process is concerned with defining a page-oriented organization (i.e. a layout) for the document content. This process can operate in two ways.

The layout process can generate a document which is not intended to be modified; it is suitable only for input to the imaging process. Such a document is said to be in formatted form.

This process can also generate a formatted processable form document which can be processed further if desired; it is suitable for input to any of the imaging, layout or editing processes.