

DIN EN 62605

ICS 35.240.30

Einsprüche bis 2018-02-22
Vorgesehen als Ersatz für
DIN EN 62605:2012-03**Entwurf**

**Multimediageräte und -systeme –
Multimedia e-publishing und e-books –
Austauschformat für elektronische Wörterbücher
(IEC 62605:2016);
Englische Fassung EN 62605:2016**

Multimedia systems and equipment –
Multimedia e-publishing and e-books –
Interchange format for e-dictionaries
(IEC 62605:2016);
English version EN 62605:2016

Systèmes et équipements multimedia –
Publication et livres électroniques multimedia –
Format d'échange pour les dictionnaires électroniques
(IEC 62605:2016);
Version anglaise EN 62605:2016

Anwendungswarnvermerk

Dieser Norm-Entwurf mit Erscheinungsdatum 2017-12-22 wird der Öffentlichkeit zur Prüfung und Stellungnahme vorgelegt.

Weil die beabsichtigte Norm von der vorliegenden Fassung abweichen kann, ist die Anwendung dieses Entwurfs besonders zu vereinbaren.

Stellungnahmen werden erbeten

- vorzugsweise online im Norm-Entwurfs-Portal von DIN unter www.din.de/go/entwuerfe bzw. für Norm-Entwürfe der DKE auch im Norm-Entwurfs-Portal der DKE unter www.entwuerfe.normenbibliothek.de, sofern dort wiedergegeben;
- oder als Datei per E-Mail an dke@vde.com möglichst in Form einer Tabelle. Die Vorlage dieser Tabelle kann im Internet unter www.din.de/go/stellungnahmen-norm-entwuerfe oder für Stellungnahmen zu Norm-Entwürfen der DKE unter www.dke.de/stellungnahme abgerufen werden;
- oder in Papierform an die DKE Deutsche Kommission Elektrotechnik Elektronik Informationstechnik in DIN und VDE, Stresemannallee 15, 60596 Frankfurt am Main.

Die Empfänger dieses Norm-Entwurfs werden gebeten, mit ihren Kommentaren jegliche relevanten Patentrechte, die sie kennen, mitzuteilen und unterstützende Dokumentationen zur Verfügung zu stellen.

Gesamtumfang 238 Seiten

DKE Deutsche Kommission Elektrotechnik Elektronik Informationstechnik in DIN und VDE

Anwendungsbeginn

Anwendungsbeginn dieser Norm ist

Nationales Vorwort

Für dieses Dokument ist das nationale Arbeitsgremium K 742 „Audio-, Video- und Multimediasysteme, -geräte und -komponenten“ der DKE Deutsche Kommission Elektrotechnik Elektronik Informationstechnik im DIN und VDE (www.dke.de) zuständig.

Die enthaltene IEC-Publikation wurde vom TC 100 „Audio, video and multimedia systems and equipment“ erarbeitet.

Das IEC-Komitee hat entschieden, dass der Inhalt dieser Publikation bis zu dem Datum (maintenance result date) unverändert bleiben soll, das auf der IEC-Website unter „<http://webstore.iec.ch>“ zu dieser Publikation angegeben ist. Zu diesem Zeitpunkt wird entsprechend der Entscheidung des Komitees die Publikation

- bestätigt,
- zurückgezogen,
- durch eine Folgeausgabe ersetzt oder
- geändert.

Für den Fall einer undatierten Verweisung im normativen Text (Verweisung auf ein Dokument ohne Angabe des Ausgabedatums und ohne Hinweis auf eine Abschnittsnummer, eine Tabelle, ein Bild usw.) bezieht sich die Verweisung auf die jeweils aktuellste Ausgabe des in Bezug genommenen Dokuments.

Für den Fall einer datierten Verweisung im normativen Text bezieht sich die Verweisung immer auf die in Bezug genommene Ausgabe des Dokuments.

Der Zusammenhang der zitierten Dokumente mit den entsprechenden deutschen Dokumenten ergibt sich, soweit ein Zusammenhang besteht, grundsätzlich über die Nummer der entsprechenden IEC-Publikation. Beispiel: IEC 60068 ist als EN 60068 als Europäische Norm durch CENELEC übernommen und als DIN EN 60068 ins Deutsche Normenwerk aufgenommen.

Das Präsidium des DIN hat mit Präsidialbeschluss 1/2004 festgelegt, dass DIN-Normen, deren Inhalt sich auf internationale Arbeitsergebnisse der Informationsverarbeitung gründet, unter bestimmten Bedingungen allein in englischer Sprache veröffentlicht werden dürfen. Diese Bedingungen sind für die vorliegende Norm erfüllt.

Da sich die Benutzer des vorliegenden Normentwurfs der englischen Sprache als Fachsprache bedienen, wird die Englische Fassung der EN 62605 veröffentlicht. Zu deren Abschnitt 1, der den Anwendungsbereich festlegt, und Abschnitt 3, der die Begriffe festlegt, wurde eine Übersetzung angefertigt und als informativer Nationaler Anhang NA des vorliegenden Normentwurfs hinzugefügt. Für die meisten der verwendeten Begriffe existieren keine gebräuchlichen deutschen Benennungen, da sich die deutschen Anwender in der Regel ebenfalls der englischen Benennungen bedienen. Dieser Normentwurf steht nicht in unmittelbarem Zusammenhang mit Rechtsvorschriften und ist nicht als Sicherheitsnorm anzusehen.

Es ist vorgesehen, auch bei der entsprechenden zukünftigen Deutschen Norm auf die deutsche Sprachfassung zu verzichten und diese in der englischsprachigen Fassung zu veröffentlichen.

Änderungen

Gegenüber DIN EN 62605:2012-03 wurden folgende Änderungen vorgenommen:

- a) Das Ref Element zur Vereinfachung der Querverweise zwischen Einträgen wurde hinzugefügt,
- b) Eine neue Version des LeXML Formats, das ein Basis-Format der ersten Ausgabe ist, wurde im Anhang B erweitert (das existierende Format wird zum Anhang A).

Nationaler Anhang NA (informativ)

Auszugsweise Übersetzung^{N1)}

1 Anwendungsbereich

In dieser Internationalen Norm wird ein Austauschformat für E-Wörterbücher für den Datenaustausch zwischen Verlegern (Publisher), Datenaufbereitern und Herstellern festgelegt.

Folgende Themen werden von dieser Internationalen Norm nicht behandelt:

- Datenformate für Lesegeräte;
- notwendige Elemente für die abschließende Druck-Reproduktion;
- Wiedergabeaspekte bezüglich physikalischer Geräte;
- Sicherheitsaspekte wie DRM für Dokumente.

3 Begriffe

Für die Anwendung dieses Dokuments gilt der folgende Begriff.

3.1

Hersteller

Organisation oder Person, die Hardware und/oder Software für das E-Book herstellt

^{N1)} Nationale Fußnote: Die Nummerierung der IEC 62605:2016-02 wurde beibehalten.

– Leerseite –

English Version

Multimedia systems and equipment - Multimedia e-publishing
and e-books - Interchange format for e-dictionaries
(IEC 62605:2016)

Systèmes et équipements multimedia - Publication et livres
électroniques multimedia - Format d'échange pour les
dictionnaires électroniques
(IEC 62605:2016)

Multimediasysteme und -systeme - Multimedia e-publishing
und e-books - Austauschformat für elektronische
Wörterbücher
(IEC 62605:2016)

This European Standard was approved by CENELEC on 2016-03-16. CENELEC members are bound to comply with the CEN/CENELEC Internal Regulations which stipulate the conditions for giving this European Standard the status of a national standard without any alteration.

Up-to-date lists and bibliographical references concerning such national standards may be obtained on application to the CEN-CENELEC Management Centre or to any CENELEC member.

This European Standard exists in three official versions (English, French, German). A version in any other language made by translation under the responsibility of a CENELEC member into its own language and notified to the CEN-CENELEC Management Centre has the same status as the official versions.

CENELEC members are the national electrotechnical committees of Austria, Belgium, Bulgaria, Croatia, Cyprus, the Czech Republic, Denmark, Estonia, Finland, Former Yugoslav Republic of Macedonia, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, the Netherlands, Norway, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, Switzerland, Turkey and the United Kingdom.



European Committee for Electrotechnical Standardization
Comité Européen de Normalisation Electrotechnique
Europäisches Komitee für Elektrotechnische Normung

CEN-CENELEC Management Centre: Avenue Marnix 17, B-1000 Brussels

European foreword

The text of document 100/2430/CDV, future edition 2 of IEC 62605, prepared by Technical Area 10 “Multimedia e-publishing and e-book technologies” of IEC/TC 100 “Audio, video and multimedia systems and equipment” was submitted to the IEC-CENELEC parallel vote and approved by CENELEC as EN 62605:2016.

The following dates are fixed:

- latest date by which the document has to be (dop) 2017-05-18
implemented at national level by
publication of an identical national
standard or by endorsement
- latest date by which the national (dow) 2019-11-18
standards conflicting with the
document have to be withdrawn

This document supersedes EN 62605:2011.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. CENELEC [and/or CEN] shall not be held responsible for identifying any or all such patent rights.

Endorsement notice

The text of the International Standard IEC 62605:2016 was approved by CENELEC as a European Standard without any modification.

In the official version, for Bibliography, the following notes have to be added for the standards indicated:

IEC 62448:2013	NOTE	Harmonized as EN 62448:2014.
ISO 3166-1	NOTE	Harmonized as EN ISO 3166-1.

CONTENTS

European Foreword	3
INTRODUCTION.....	6
1 Scope.....	7
2 Normative references	7
3 Terms and definitions	7
4 Position and requirements for interchange format for e-dictionaries	7
4.1 Interchange format for e-dictionaries in contents creation/distribution model.....	7
4.2 Requirements for interchange format for e-dictionaries	8
5 File formats	8
6 Semantics	9
Annex A (normative) XMDF-LeXML format	10
A.1 General.....	10
A.2 Overview of the format's structure	10
A.3 Elements and attributes.....	11
A.3.1 General	11
A.3.2 Page_ID	11
A.3.3 Object_ID	11
A.3.4 Char_ID	11
A.3.5 Reading	11
A.3.6 Filename.....	12
A.3.7 Standard character	12
A.3.8 Standard character string	12
A.3.9 Extended character	13
A.3.10 Extended character string.....	13
A.3.11 External character.....	13
A.3.12 External character string	15
A.3.13 External extended character string	15
A.3.14 Coordinates	15
A.3.15 Polygonal_region	16
A.3.16 Color	16
A.3.17 Date	17
A.3.18 Time	17
A.3.19 Country.....	17
A.3.20 Personal_name	17
A.3.21 Organization_name	19
A.3.22 Address	19
A.3.23 Permission.....	21
A.3.24 Keyword	23
A.3.25 Telephone_number	23
A.3.26 Mail_address	24
A.4 Description format details	25
A.4.1 General	25
A.4.2 Book information modules <bvf>.....	25
A.4.3 Content management module <body_module>	40

A.4.4	Event info module <event_info>	58
A.4.5	Parts data module <parts_module>	65
A.4.6	Object instances	71
A.5	Available color names	118
A.6	Localization	119
A.6.1	Possible additions.....	119
A.6.2	Standard characters.....	119
A.6.3	Characters usable for reading	120
A.6.4	Line breaking methods.....	120
A.6.5	Sorting rules for <search_table_def>	121
A.6.6	Additional attributes for <enable_key_type>	122
A.6.7	Normalization methods for <key_normalization>	124
A.6.8	Character encoding conversions	124
A.7	Adaptation	124
A.8	Specification of the XMDF-LeXML format in the RELAX NG syntax	125
Annex B (normative)	LeXML format.....	161
B.1	General	161
B.2	Elements for content structure.....	161
B.2.1	Parameter entity definition	161
B.2.2	Root elements	163
B.2.3	headword-related elements	167
B.2.4	Main-text related elements	170
B.2.5	Subheadword related elements	174
B.2.6	Other block elements	179
B.2.7	Media related elements	183
B.2.8	Other structural elements	185
B.3	Inline elements	190
B.3.1	Labels	190
B.3.2	pronunciation/accent – related elements.....	192
B.3.3	Part-of-speech related and other elements	194
B.3.4	Other dictionary-specific elements.....	195
B.3.5	Text-decoration related elements	205
B.3.6	Typesetting-related elements	210
B.3.7	Other elements	212
B.4	Specification of the LeXML format in the DTD syntax.....	219
Bibliography		233
Annex ZA (normative)	Normative references to international publications with their corresponding European publications.....	236
Figure 1	– Contents creation/distribution model.....	7
Figure 2	– Contents creation/distribution model (modified).....	8
Figure 3	– Relationship between concepts	8
Figure A.1	– XML tree structure	10
Figure A.2	– Example of valign="middle"	73
Figure A.3	– Example of dropped capital	79
Figure A.4	– Left and right margin of a paragraph.....	79

Figure A.5 – Horizontal writing in vertical text	87
Figure A.6 – Ruby	88
Figure A.7 – Example of search page object instance rendering	114
Table A.1 – Base characters for reading	12
Table A.2 – Standard character set	12
Table A.3 – Usable characters for a telephone number	24
Table A.4 – Characters usable for email addresses	24
Table A.5 – Characters usable for the lookup key	57
Table A.6 – Color names	119
Table A.7 – Examples of additional standard character sets	120
Table A.8 – Example of additional characters usable for readings	120
Table A.9 – Example of additional sorting rules	122
Table A.10 – Example of additional language specific attributes for <enable_key_type>	123

INTRODUCTION

Markets for multimedia e-books and e-publishing require standardization of formats for e-book data interchange among associated people, authors, data preparers, publishers and readers. The formats are classified into submission format, interchange format and reader's format. The submission format supports an interaction between authors and data preparers. The reader's format depends on e-publishing equipment. The interchange format provides an interchange format for data preparers and publishers and therefore should be e-publishing equipment independent.

The International Electrotechnical Commission (IEC) draws attention to the fact that it is claimed that compliance with this document may involve the use of patents.

IEC takes no position concerning the evidence, validity and scope of this patent right.

The holder of this patent right has assured the IEC that he/she is willing to negotiate licences under reasonable and non-discriminatory terms and conditions with applicants throughout the world. In this respect, the statement of the holder of this patent right is registered with IEC. Information may be obtained from:

Sharp Corporation,
22-22 Nagaike-cho,
Abeno-ku,
Osaka 545-8522,
Japan

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights other than those identified above. IEC shall not be held responsible for identifying any or all such patent rights.

ISO (www.iso.org/patents) and IEC (<http://patents.iec.ch>) maintain on-line data bases of patents relevant to their standards. Users are encouraged to consult the data bases for the most up to date information concerning patents.

MULTIMEDIA SYSTEMS AND EQUIPMENT – MULTIMEDIA E-PUBLISHING AND E-BOOKS – INTERCHANGE FORMAT FOR E-DICTIONARIES

1 Scope

This International Standard specifies the interchange format for e-dictionaries among publishers, content creators and manufacturers.

This International Standard does not address the following aspects:

- data formats for reading devices;
- elements necessary for final print reproduction only;
- rendering issues related to physical devices;
- security issues such as DRM for documents.

2 Normative references

The following documents, in whole or in part, are normatively referenced in this document and are indispensable for its application. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

IEC TS 62229:2006, *Multimedia systems and equipment – Multimedia e-publishing and e-book – Conceptual model for multimedia e-publishing*

3 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

3.1

manufacturer

organization or person that manufactures hardware and/or software of the e-book

4 Position and requirements for interchange format for e-dictionaries

4.1 Interchange format for e-dictionaries in contents creation/distribution model

The conceptual model for multimedia e-publishing (IEC TS 62229) defines a contents creation/distribution model shown in Figure 1.

Author <--(1)--> Data preparer <--(2)--> Publisher --(3)--> Reader
IEC

Key

- (1) content data in submission format
- (2) content data in interchange format
- (3) content data in reader's format

Figure 1 – Contents creation/distribution model

It should be noted that the role of manufacturers of e-dictionary hardware and software overlaps that of the publisher in Figure 1. Therefore, a slightly modified model will be assumed for this International Standard, as shown in Figure 2.

Author <--(1)--> Data preparer <--(2)--> Publisher (manufacturer) --(3)--> Reader^{IEC}

Figure 2 – Contents creation/distribution model (modified)

This International Standard specifies the interchange format between data preparers and publishers, i.e. a format for (2) in Figure 2, though it may be used as a reader's format.

4.2 Requirements for interchange format for e-dictionaries

An interchange format for e-dictionaries needs to address the following.

- Description of keywords, links from the keywords to entries (link data) and the order of the entries.
- Description of articles for each entry (entry data). This includes text, image, and other multimedia functionalities generally required for e-books.
- Description of bibliographical data and other data. This should include the name of the author and the publisher, the title of the content and the explanatory note. The relationship between these concepts is visually represented in Figure 3.
- Description of contents written in various languages.

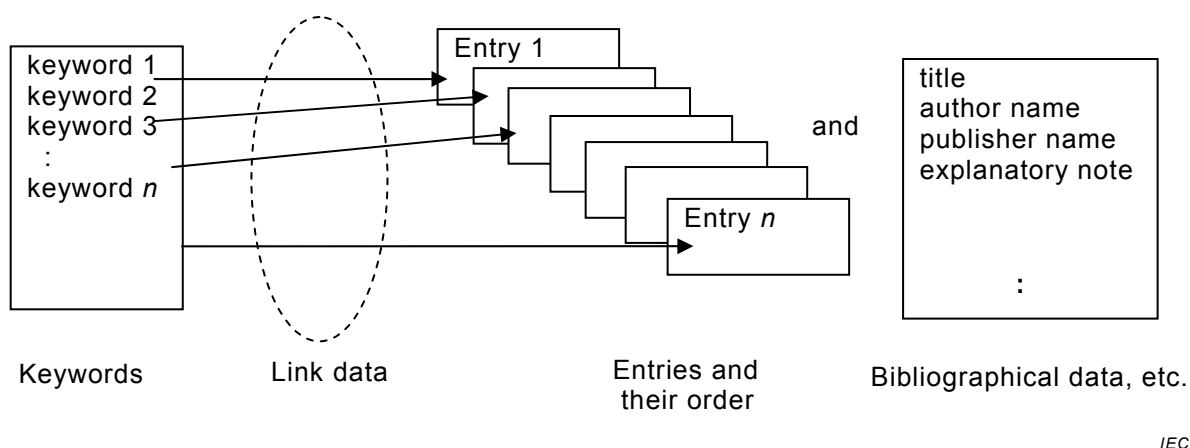


Figure 3 – Relationship between concepts

5 File formats

This International Standard defines two XML-based formats. One is based on XMDF (as described in IEC 62448:2013, Annex B) and LeXML. The format is hereafter called XMDF-LeXML format. The other is based solely on LeXML 3.0. They are presented in Annex A and Annex B, respectively.

NOTE LeXML is proposed by Digital ASSIST Ltd. Its original specifications are found at <http://www.d-assist.com/index.html> (in Japanese).

6 Semantics

Elements of the XMDF-LeXML format can be rendered in accordance with appropriate style specifications, which are outside the scope of this International Standard.

Annex A (normative)

X MDF-LeXML format

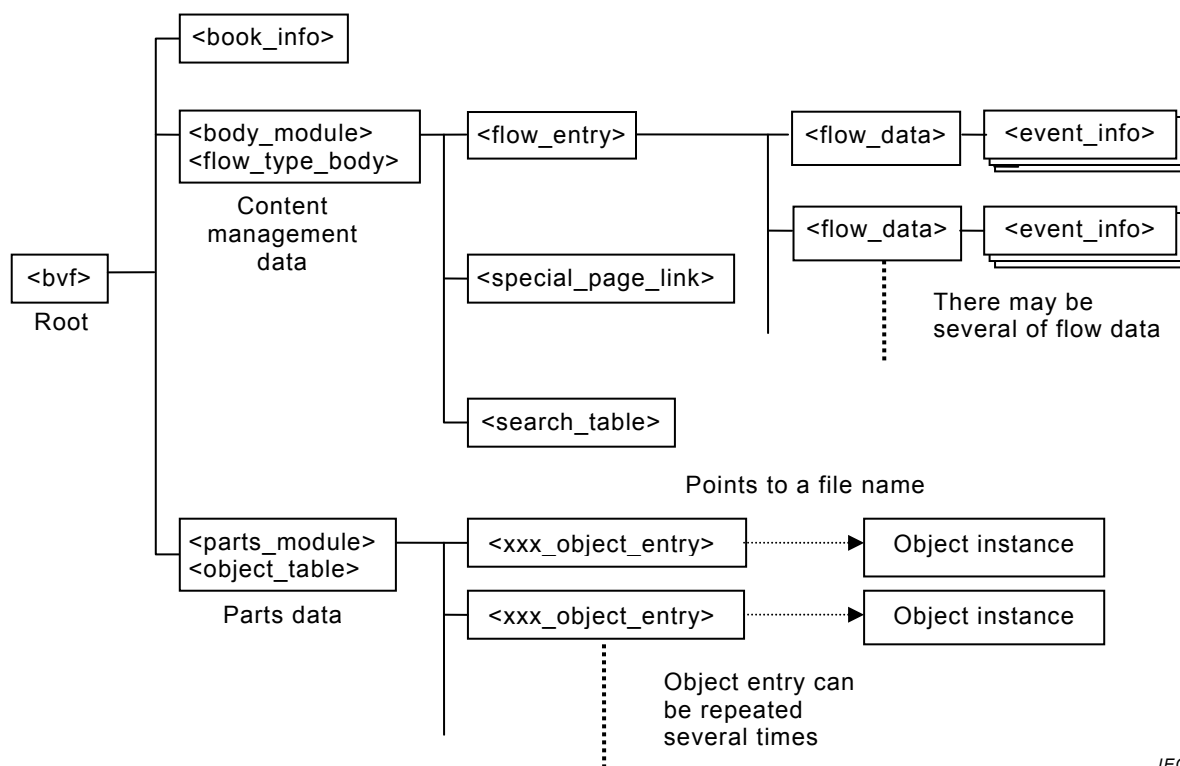
A.1 General

The X MDF-LeXML format is an interchange format for e-dictionaries multimedia e-book data interchange, targeted at data preparers and publishers rather than the reader, with an emphasis on mobile devices as a target platform. Much like HTML, this format does not split the document in fixed pages, but determines the layout according to the viewer device's display size, the font in use, and so on. In this annex, such contents will be referred to as flowing content, as opposed to paged content.

A.2 Overview of the format's structure

Flowing contents are usually composed of several concatenated flows. This annex makes no particular requirement concerning the way the flowing content should be split into individual flows. This decision is left to the data preparer, to accommodate the various types of contents. For instance, a newspaper may have one flow per article, a novel one per chapter, and so on. It is also possible not to split the content, and to have only one flow. However, it should be noted that particularly large flows, or an extremely large number of flows, may impact on runtime performance, depending on the specific version of the viewer in use, the available memory, and so on.

The XML tree structure of the format is shown in Figure A.1.



IEC

Figure A.1 – XML tree structure

The actual contents of each flow, in other words, what will be displayed by the viewer, is recorded in the object instance. The object instance is registered in object_entry, and associated with an ID number and other auxiliary data, turning it into playable/displayable data. Flow_data determines its content by pointing at such registered objects. In addition, information on functionalities such as page link is recorded in event_info.

The main part of Annex A is generic, and may be used for any country and language. However, some parts may have language specific behavior. Localization-related issues are detailed in Clause A.6.

A.3 Elements and attributes

A.3.1 General

The different types of values that may be used in the various elements or attributes are explained below. The elements and attributes detailed below will be valid throughout this annex, and will be referenced by other constructs. In the following explanations, alphanumeric characters refer to numerals from 0 to 9 and alphabetic letters from a to z and A to Z.

A.3.2 Page_ID

Page_ID specifies a unique identification number for the flow data of the flowing contents. It is a string starting by the "PG" characters, followed by alphanumeric characters.

Example:

```
<flow_data flow_id="PG0002" ... />
```

A.3.3 Object_ID

Object_ID specifies a unique identification number for objects used in the flowing contents. It is a string starting by the "OB" characters, followed by alphanumeric characters.

Example:

```
<dynamic_text_object_entry id="OB0ue4".../>
```

A.3.4 Char_ID

Char_ID specifies an identification number for positions (character strings, etc.) within text and dictionary data objects. It is an alphanumeric string which is to be given uniquely in the text (see A.4.6.2) and dictionary data object instance (see A.4.6.3). Char IDs with the same value in different object instances are regarded as separate and don't affect each other.

Example:

```
<trigger_pointer id="OB29s0/CR0de4"/>  
Click<char_id char_id="CR0001">here</char_id>for details.
```

A.3.5 Reading

For sorting purposes, reading may be useful to specify the reading of each word. Restricting the characters allowed for this purpose to a limited set makes it easier to define the sorting method.

Such characters should be determined on a per language basis. All languages can use the characters listed in Table A.1 as a common base, while the localization (see Clause A.6) will describe the language specific extensions to it.

Table A.1 – Base characters for reading

Name	Corresponding characters ^a
Basic alphabet	A to Z (U+0041 to U+005A) A to z (U+0061 to U+007A)
Numerals	0 to 9 (U+0030 to U+0039)
Others	space (U+0020), ((U+0028),) (U+0029)
^a [Attributes]ll values are in Unicode.	

Example:

```
<title reading="PI">π</title>
```

A.3.6 Filename

Filenames should be written using the following convention. The path is relative to the file in which this reference is made. Network paths should not be used. For portability concerns, it is recommended that only ACSII characters be used. Both the slash and backslash characters are acceptable as directory separators. It is also recommended not to use excessively long filenames, as those might not be supported by the host operating system.

Example:

```
<dynamic_text_object_entry src="sect1.xml" type="text/x-bvf-text" id="OB0ue4"/>
```

A.3.7 Standard character

The standard character set of the document, as set by the default_ccs attribute of the <bvf> element (see A.4.2), is to be chosen from a well defined list, so as to ease the development of viewing software. However, this list may change for different localized versions of the XMDF-LeXML format. Any e-book data has to define its standard character set as one of or a combination of character set(s) listed in Table A.2 and those defined for a specific localization (see A.6.2).

Table A.2 – Standard character set

Character set name	Description
"ISO 646-IRV"	Characters in the range of US-ASCII

A.3.8 Standard character string

A string composed of Standard characters is called a Standard character string. Unless specified otherwise, the spacing characters (space (U+0020), linefeed (U+000D, U+000A, U+000D+U+000A), tabulation (U+0009)) are to be handled as follows:

Space (U+0020) to be displayed as is.

Linefeed (U+000D, U+000a, U+000D+U+000A) not to be displayed, but simply ignored.

Tabulation (U+0009) to be displayed as if it were a single space.

Furthermore, because of restriction in the XML format, linefeeds (U+000D, U+000A, U+000D+U+000A) and tabulations (U+0009) in attribute values should be replaced by spaces when converting to the distribution format.

A.3.9 Extended character

Characters which have Unicode code points while not being among those listed below are referred to as Extended characters.

Standard characters

Surrogate pair range (U+D800 to U+DFFF)

BOM (Byte Order Mark) (U+FFFE,U+FEFF)

NON CHARACTER (U+FFFF)

Control characters (characters between U+0000 and U+001F except tabulation (U+0009) and linefeed (U+000A, U+000D), as well as DEL (U+007F)).

If an e-book indeed uses any Extended character in its data, the name of a character set that covers those Extended characters should be appended to the default_ccs attribute of the <bvf> element. Note that all Extended characters used in the document do not need to be covered by the same character set, as it is possible to specify several character sets.

A.3.10 Extended character string

A string composed of Standard characters and Extended characters is called an Extended character string. Unless specified otherwise, the spacing characters (space (U+0020), linefeed (U+000D, U+000A, U+000D+U+000A), tabulation (U+0009)) are to be handled the same way as in Standard character strings.

A.3.11 External character

To display a character which is neither a Standard character nor an Extended character, it is possible to use the <external_char> element described below.

<external_char> inserts an External character. The viewer may display it according to the following methods.

- a) Display the character set by the alt_set and alt_code attributes.
- b) Display the image set by the alt_img or alt_vimg attributes.
- c) Display the alternative letter set by the alt attribute.

Its syntax is given in Relax NG compact format below and explained in the following text.

NOTE For definitions that appear in the Relax NG compact representation of each element in this annex, see Clause A.8.

external_char = element external_char { attlist_external_char, text }

attlist_external_char &=

attribute alt_set { text }?,

attribute alt_code { text }?,

attribute alt_img { text }?,

attribute alt_vimg { text }?,

attribute img_type { text }?,

attribute alt { text }?

[Attributes]

alt_set: Together with the alt_code attribute, it allows designating the External character to be used. This alt_set attribute indicates the font name, while the alt_code attribute indicates the character code point within the font. The alt_set attribute is written in the following way:

alt_set = "font1,font2, ..."

The alt_set attribute may hold several font names, separated by "," (U+002C). In that case, the viewer should use the first font of the list that is available (either from the platform, or included in the contents data itself) to display the character.

alt_code: Selects a character code point in the font specified by the alt_set attribute. It may be written either as a decimal number or a hexadecimal number, prefixed by "0x". In case several fonts have been defined in the alt_set attribute, the character code shall represent the same character in all of them. This attribute can be omitted.

alt_img: Defines an alternative character image. Written as a Filename. Before opening the file indicated by this attribute, the img_type attribute should be checked for authorized file types. Note that it may be used only when <external_char> is used in a text object instance. When both alt_img and alt_vimg are used, the file types shall match. This can be omitted. When this attribute is set, display should be done according to the following methods:

a) Monochrome images

Black pixels represent the letter, and white pixels the background. The font color and background color are to be displayed according to the color attribute of the element.

b) Images with levels of gray

Black pixels represent the letter, and white pixels the background. The font color and background color are to be displayed according to the color attribute of the element. The color of "gray" pixels shall be computed as an intermediate value between the font color and the background color.

c) Color images

Displayed as is.

alt_vimg: Defines an alternative character image to be used when the text is displayed vertically (as can be the case in some languages, such as Japanese). Written as a Filename. When omitted, the image defined in the alt_img attribute should be used both for horizontal and vertical layout. Before opening the file indicated by this attribute, the img_type attribute should be checked for authorized file types. Note that it may be used only when <external_char> is used in a text object instance. When both alt_img and alt_vimg are used, the file types shall match. This can be omitted. When this attribute is set, the display should be done according to the same methods as with alt_img.

img_type: Defines the mime type of the images files set in the alt_img and alt_vimg attributes. Currently, only PNG and JPEG are supported, and should be written as:

"image/png"

"image/jpeg"

When either alt_img, alt_vimg or both are set, this attribute is required. As with these two attributes, it can only be used in an <external_char> element inside a text object instance.

alt: Alternative character string. Written as a Standard character string. May be omitted.

Example:

```
<external_char alt_set="oooextchars" alt_code="47268" alt_img="ou.img"
      alt_vimg="ou_v.img" img_type="image/jpeg" alt="鷗"/>
<external_char alt="間"/>
<external_char alt_set="sharp_extchars" alt_code="0x2345" alt="間"/>
```

A.3.12 External character string

An External character string is a string composed of Standard characters, External characters, or both. Unless specified otherwise, spacing characters (space (U+0020), linefeed (U+000D, U+000A), tabulation (U+0009)) should be handled the same way as they are handled in Standard character strings.

Example:

```
森<external_char alt_img="ou.png" alt_type="image/png" alt="鷗"/>外
内田百<external_char alt_set="sharp_extchars" alt_code="0x2345" alt="間"/>
```

A.3.13 External extended character string

An External extended character string is a string of Standard characters, Extended characters, External characters, or a combination of any of the above. Unless specified otherwise, spacing characters (space (U+0020), linefeed (U+000D, U+000A), tabulation (U+0009)) should be handled the same way as they are handled in Standard character strings.

A.3.14 Coordinates

Data type to be used to store coordinates, dimension and other similar information composed of an x and a y value. It is written as "(x, y)". The name of the attribute which uses this type of data depends on the element.

The coordinate system explained below will be used in this annex. The origin is at the top left corner, the x axis oriented rightwards, and the y axis downwards. As the system of coordinates used by the viewer to map things on the screen is implementation-dependent, it will not be discussed here.

Local coordinate system:

The coordinate system local to an object takes its origin in the upper left corner of the circumscribed rectangle, and has the same orientation as the general coordinate system. Positions within an object should be expressed in the local coordinate system.

Example:

```
<vertex position="(100,200)"/>
```

A.3.15 Polygonal_region

Data format to store the apexes of a polygon, or any other ordered sequence of vertexes. Each vertex is stored in a <vertex> element. When defining the shape of a polygon, the edges shall not cross. If they do, the viewer's behavior is unspecified. Its syntax is given in Relax NG compact format below and explained in the following text.

vertex = element vertex { attlist_vertex, empty }

attlist_vertex &= attribute position { text }

[Attribute]

position: the position of the apex, expressed as Coordinates. This attribute shall not be omitted.

Example:

```
<vertex position="(100,0)"/>  <!-- in the case of a triangle -->
<vertex position="(0,100)"/>
<vertex position="(200,100)"/>
```

A.3.16 Color

Data type to define colors. The following attributes are defined.

[Attributes]

color_space: Specifies the color space to be used. Currently, only RGB is accepted. If this attribute is omitted, the viewer should act as if RGB was set.

color: Specifies the color name. Color names or numerical values may be used. Acceptable color names are listed in Table A.6. The default value depends on the actual element and context. Numerical values are to be written in the following syntax.

In RGB: written as #RRGGBB. With RR, GG, BB being hexadecimal numbers, ranging from 00 to FF. Grayscale pixel values are represented by setting RR, GG and BB to the same value.

opacity: level of opacity. Ranging from 0 (transparent) to 100 (opaque). Presently, the only admitted value is 100, and in case the attribute is omitted, it defaults to 100.

Example:

```
<font color="#FF0000"/>
      <!-- the color_space is unspecified, and thus defaults to RGB-->
<font color="#FF0000" opacity="100"/>
<font color="black"/>
```

A.3.17 Date

Data format to store dates. It uses the same representation as ISO 8601. For instance, 1994-11-05T08:15:30-05:00 corresponds to November 5, 1994, 8:15:30 am, US Eastern Standard Time. Abbreviated forms are also accepted. Please refer to <http://www.w3.org/TR/NOTE-datetime> for details.

Example:

```
<publication_date type="publish">1994</publication_date>
```

A.3.18 Time

Data format to specify durations. Written as "XXdXXhXXmXXsXXXms", where X stands for a digit between 0 and 9. For instance, "10d5h30m10s015ms" would mean 10 days, 5 hours, 30 minutes, 10 seconds and 15 milliseconds. Abbreviated forms such as "5m30ms" or "1s" are possible. There is no upper bound to the number of days.

Example:

```
<flip_animation renewal_time="1s" >
```

A.3.19 Country

Data format to specify a country name. Written according to the ISO 3166-1 alpha 3 standard, in lower case.

Example:

```
<publication_place>jpn</publication_place>
```

A.3.20 Personal_name

Data format to store people's names, such as the document author. It is stored under the <personal_name> element. Its syntax is given in Relax NG compact format below and explained in the following text. Several child elements are used to define the various parts of the name: first name, middle name and last name. This information shall be entered via the elements defined below. At least one of <first_name>, <middle_name>, and <last_name> shall be specified.

personal_name =

element personal_name {

attlist_personal_name,

((first_name?, middle_name?, last_name?)

| (first_name?, last_name?, middle_name?)

| (last_name?, first_name?, middle_name?)

| (last_name?, middle_name?, first_name?))

}

attlist_personal_name &= empty

[Child elements]

<first_name> Sets the first name. Written as an External character string. May be omitted. Its syntax is given in Relax NG compact format below and explained in the following text.

first_name = element first_name { attlist_first_name, TextWithGaiji }
attlist_first_name &= attribute reading { text }?

[Attribute]

reading: Pronunciation of the first name, written as a Reading. May be omitted.

<middle_name> Sets the middle name. Written as an External character string. May be omitted. Its syntax is given in Relax NG compact format below and explained in the following text.

middle_name = element middle_name { attlist_middle_name, TextWithGaiji }
attlist_middle_name &= attribute reading { text }?

[Attribute]

reading: Pronunciation of the middle name, written as a Reading. May be omitted.

<last_name> Sets the last name. Written as an External character string. May be omitted. Its syntax is given in Relax NG compact format below and explained in the following text.

last_name = element last_name { attlist_last_name, TextWithGaiji }
attlist_last_name &= attribute reading { text }?

[Attribute]

reading: Pronunciation of the last name, written as a Reading. May be omitted.

Example:

```
<personal_name>  
  <!-- to record John Smith -->  
  <last_name>Smith</last_name>  
  <first_name>John</first_name>  
</personal_name>
```

A.3.21 Organization_name

Data format to define company's name, such as publishers. It is stored in the <organization_name> element. Written as an External character string. Its syntax is given in Relax NG compact format below and explained in the following text.

organization_name =

element organization_name { attlist_organization_name, TextWithGaiji }

attlist_organization_name &= attribute reading { text }?

[Attribute]

reading: Pronunciation of the company's name, written as a Reading. May be omitted.

Example:

```
<organization_name >ABCD Corporation</organization_name>
```

A.3.22 Address

Data format used to define an address, telephone number, email, and other information. Address-related information is all stored in child elements of a main <address_info> element. Its syntax is given in Relax NG compact format below and explained in the following text.

address_info =

element address_info {

attlist_address_info,

postal_code?,

address?,

telephone?,

fax?,

```
mail_address?,  
  
website?,  
  
address_other_info?  
  
}  
  
attlist_address_info &= empty
```

[Child elements]

<postal_code> Stores the postal code (zip code) as a Standard character string. May be omitted. Its syntax is given in Relax NG compact format below.

```
postal_code = element postal_code { attlist_postal_code, text }  
attlist_postal_code &= empty
```

<address> Stores the address as an External character string. May be omitted. Its syntax is given in Relax NG compact format below.

```
address = element address { attlist_address, TextWithGaiji }  
attlist_address &= empty
```

<telephone> Stores the phone number as a Standard character string. May be omitted. Its syntax is given in Relax NG compact format below.

```
telephone = element telephone { attlist_telephone, text }  
attlist_telephone &= empty
```

<fax> Stores the fax number as a Standard character string. May be omitted. Its syntax is given in Relax NG compact format below.

```
fax = element fax { attlist_fax, text }  
attlist_fax &= empty
```

<mail_address> Stores the mail address as a Standard character string. May be omitted. Its syntax is given in Relax NG compact format below.

```
mail_address = element mail_address { attlist_mail_address, text }  
attlist_mail_address &= empty
```

<website> Stores the home page's URI as a Standard character string. May be omitted. Its syntax is given in Relax NG compact format below.

```
website = element website { attlist_website, text }
attlist_website &= empty
```

`<address_other_info>` Allows storing additional information not covered by the preceding elements as an External character string. May be omitted. If this information is to be displayed, spaces (U+0020) and linefeeds (U+000D, U+000A, U+000D+U+000A) should be displayed as is, while tabulations (U+0009) should be displayed as spaces. Its syntax is given in Relax NG compact format below.

```
address_other_info =
  element address_other_info {
    attlist_address_other_info, TextWithGaiji
  }
attlist_address_other_info &=
  [ a:defaultValue = "preserve" ]
attribute xml:space { "default" | "preserve" }?
```

Example:

```
<address_info>
  <postal_code>75008</postal_code>
  <address>xxx avenue des champs elysees, Paris, France </address>
  <telephone>01 53 xx xx xx</telephone>
  <mail_address>xxx@XXXXXX.fr</mail_address>
  <website>http://www.XXXXXX.fr</website>
</address_info>
```

A.3.23 Permission

Sets the permissions, such as the right to print, or copy. The various permissions are stored in child elements of the `<permission_info>` element. When the `<permission_info>` element is omitted, all permissions are set to the same value as when each permission element is omitted. The following explanations refer to an "authenticated user". Under usual circumstances, all viewers of the document are considered authenticated. However, the distribution format may include DRM technologies and authentication mechanisms. The syntax of the `<permission_info>` element is given in Relax NG compact format below and explained in the following text.

```
permission_info =
  element permission_info {
    attlist_permission_info, print_permission?, copy_permission?
  }
attlist_permission_info &= empty
```

[Child elements]

<print_permission> Defines whether printing is permitted or not. When omitted, it is to be considered as if the permission attribute was set to "no". Its syntax is given in Relax NG compact format below and explained in the following text.

```
print_permission =  
  element print_permission { attlist_print_permission, empty }  
attlist_print_permission &=  
  [ a:defaultValue = "no" ]  
  attribute permission { "authorized" | "no" }?
```

[Attribute]

permission: Defines whether printing is permitted or not. The following values are possible.

"authorized": authenticated users may print.

"no": no one can print (default value).

<copy_permission> Define whether copying is permitted or not. When omitted, it is to be considered as if the permission attribute was set to no. Its syntax is given in Relax NG compact format below and explained in the following text.

```
copy_permission =  
  element copy_permission { attlist_copy_permission, empty }  
attlist_copy_permission &=  
  [ a:defaultValue = "no" ]  
  attribute permission { "authorized" | "in_device_only" | "no" }?
```

[Attribute]

permission: Defines whether copying is permitted or not. The following values are possible.

"authorized": authenticated users may copy.

"in_device_only": authenticated users may copy, but only within the viewer device. If the device does not provide mechanisms to prevent external copy, then copy is forbidden.

"no": No one can copy (default value).

Example:

```
<permission_info>  
  <print_permission permission="authorized"/>  
</permission_info>
```

A.3.24 Keyword

With the <keyword_list> element, it is possible to attach a list of keywords to the bibliographical data, to a flow data, or to an object (in the present standard, it is limited to bibliographical data). A <keyword_list> should contain one or more <keyword> elements as child element, each of these recording one keyword. The syntax of the <keyword> element is given in Relax NG compact format below and explained in the following text.

```
keyword = element keyword { attlist_keyword, TextWithGaiji }
```

```
attlist_keyword &=
```

```
  attribute category { text }?,
```

```
  attribute reading { text }?
```

[Attributes]

category: Defines the category the keyword belongs to, written as a Standard character string. May be omitted.

reading: Records the pronunciation of the keyword, written as a Reading. May be omitted.

[Child elements]

External character string: Records the actual keyword.

Example:

```
<keyword_list>
  <keyword category="History" >Renaissance</keyword>
  <keyword>xml</keyword>
</keyword_list>
```

A.3.25 Telephone_number

Data format to record telephone numbers. It allows dialing such a number when the viewer device is a telephone. It is written as a combination of the characters listed in Table A.3. The number of characters shall be between 1 and 64, inclusive.

If a function defined by the phone number cannot be executed, the viewer should not make the call. Namely,

when the phone number string is too long for the device to handle,

when the phone number string includes characters not listed in Table A.3,

when the phone number string contains a function which cannot be executed.

Table A.3 – Usable characters for a telephone number

Character	ASCII code	Meaning
0 to 9	0x30 to 0x39	number
#	0x23	# button
*	0x2A	* button
-	0x2D	Ignored.
,	0x2C	Pause (one second). If a one second pause cannot be made, wait for key press instead.
/	0x2F	Pause (wait for key press).
P or p	0x50 or 0x70	Pause (wait for signal). If it is not possible to wait for the signal, wait for a key press instead.
+	0x2B	Sign to make an international call. Only at the beginning of a phone number. (If entered in another position, do not make the call.)

A.3.26 Mail_address

Data format to store an email address. Written as local-part@domain. local-part and domain may use any of the characters recorded in Table A.4. The maximum length for local-part@domain is 256 bytes. Within the following characters, "&" (0x26) shall be written as an xml entity: "&". Such xml entities are counted as one byte.

Table A.4 – Characters usable for email addresses

Category	Characters	ASCII code
numerals	0 to 9	0x30 to 0x39
alphabet	A to Z	0x41 to 0x5A
	a to z	0x61 to 0x7A
!	!	0x21
\$	\$	0x24
%	%	0x25
&	&	0x26
*	*	0x2A
+	+	0x2B
-	-	0x2D
.	.	0x2E
/	/	0x2F
=	=	0x3D
?	?	0x3F
^	^	0x5E
_	_	0x5F
~	~	0x7E

A.4 Description format details

A.4.1 General

Each book takes the form of an xml document as can be seen below.

```
<?xml version="1.0" encoding="UTF-8" ?>
<bvf id="1234" id-type="...">
    <book_info>
        ...      <!-- Bibliography comes here -->
    </book_info>

    <body_module>
        ...      <!--Flow data comes here -->
    </body_module>

    <parts_module>
        ...      <!-- Objects are registered here -->
    </parts_module>
</bvf>
```

The character encoding used in the document is specified in the usual XML way, with the encoding attribute of the <xml> element. This annex recommends using UTF-8 or UTF-16 to avoid conversion problems, see A.6.8 for further details.

A.4.2 Book information modules <bvf>

A.4.2.1 General

The book information module is recorded in the <bvf> element. It serves as a root element for all data in the book, all information related to the document are stored inside it (it may also happen that only the filename of external files is stored here). The syntax of the <bvf> element is given in Relax NG compact format below and explained in the following text.

bvf = element bvf { attlist_bvf, book_info, body_module, parts_module }

attlist_bvf &=

attribute id_type { text }?,

attribute id { text }?,

attribute default_ccs { text },

attribute display_size { text }?

[Attributes]

id_type: Defines what type of number is stored in the id attribute. Written as a Standard character string. May be omitted.

- id: Records the Identification number of this book, in the system specified by the id_type attribute. Written as a Standard character string. May be omitted. Not to be confused with Reference ID (see A.4.6.3.4).
- default_ccs: Sets the name of the character set of the Standard characters (see A.3.7) and Extended characters used in this annex (see A.3.9). When more than one character set is specified, they are separated by a "," (U+002C). Shall not be omitted.
- display_size: Specifies the display (screen) size that was assumed while creating the contents, written in the standard coordinates format. May be omitted.

[Child elements]

- <book_info> Records bibliographical data. Shall not be omitted. See A.4.2.2 for details.
- <body_module> Content management module. Shall not be omitted. See A.4.3 for details.
- <parts_module> Parts data modules. Shall not be omitted. See A.4.5 for details.

Example:

```
<bvf id_type="ISBN" id="xxx-x-xxxx-xxxx-x" default_ccs="ISO 646-IRV">
    <book_info> ... </book_info>      <!-- Bibliography -->
    <body_module> ... </body_module>  <!-- Content module -->
    <parts_module> ... </parts_module> <!-- Parts data modules -->
</bvf>
```

A.4.2.2 Bibliographical data <book_info>

This is where bibliographical data, such as the author or the title, is stored. The syntax of the <book_info> element is given in Relax NG compact format below and explained in the following text.

book_info =

```
element book_info {
    attlist_book_info,
    title_info,
    author_info?,
    publisher_info?,
    seller_info?,
    book_id_info?,
    classification_info?,
    rating?,
```

```
    publication_place?,  
    publication_date_info?,  
    net_price_info?,  
    book_abstract?,  
    front_cover_image?,  
    spine_cover_image?,  
    keyword_list?,  
    other_book_info?  
}  
  
attlist_book_info &= empty  
  
title_info =  
  
    element title_info {  
        attlist_title_info,  
        series_title?,  
        title,  
        subtitle?,  
        edition_info?,  
        title_other_info?  
    }  
}
```

[Child elements]

<title_info> Stores the information related to the title. Shall not be omitted. Its syntax is given in Relax NG compact format below and explained in the following text.

```
title_info =  
    element title_info {  
        attlist_title_info,  
        series_title?,  
        title,  
        subtitle?,
```

```

edition_info?,
title_other_info?
}
attlist_title_info &= empty

```

[Child elements]

<series_title> The title of the series is recorded as an External character string in this element. If there is no series' title, it may be omitted. Its syntax is given in Relax NG compact format below and explained in the following text.

```

series_title =
  element series_title { attlist_series_title, TextWithGaiji }
attlist_series_title &= attribute reading { text }?

```

[Attribute]

reading: gives the pronunciation of the series' title as a Reading. May be omitted.

<title> The title is recorded as an External character string in this element. It shall not be omitted. Its syntax is given in Relax NG compact format below and explained in the following text.

```

title = element title { attlist_title, TextWithGaiji }
attlist_title &= attribute reading { text }?

```

[Attribute]

reading: gives the pronunciation of the title as a Reading. May be omitted.

<subtitle> The subtitle is recorded as an External character string in this element. If there is no subtitle, it may be omitted. Its syntax is given in Relax NG compact format below and explained in the following text.

```

subtitle = element subtitle { attlist_subtitle, TextWithGaiji }
attlist_subtitle &= attribute reading { text }?

```

[Attribute]

reading: gives the pronunciation of the subtitle as a Reading. May be omitted.

<edition_info> Information concerning the revision history of the book is recorded as an External character string in this element. May be

omitted. If the information is to be displayed, the space character (U+0020), and the linefeed and carriage return characters (U+000D, U+000A, U+000D+U+000A) are to be displayed as is, but the tabulation character (U+0009) is to be displayed as if it were a space. Its syntax is given in Relax NG compact format below and explained in the following text.

```
edition_info =
  element edition_info { attlist_edition_info, TextWithGaiji }
attlist_edition_info &=
  attribute this_version { text }?,
  [ a:defaultValue = "preserve" ]
  attribute xml:space { "default" | "preserve" }?
```

[Attribute]

this_version: Specifies which of the cited versions is the present one. For instance, "Third revision". May be omitted.

<title_other_info> Other information related to the title may be stored as an External character string in this element. May be omitted. If the information is to be displayed, the space character (U+0020), and the linefeed and carriage return characters (U+000D, U+000A, U+000D+U+000A) are to be displayed as is, but the tabulation character (U+0009) is to be displayed as if it were a space. Its syntax is given in Relax NG compact format below.

```
attlist_title_other_info &=
  [ a:defaultValue = "preserve" ]
  attribute xml:space { "default" | "preserve" }?
```

<author_info> Stores the information related to the author. May be omitted. Its syntax is given in Relax NG compact format below and explained in the following text.

```
author_info = element author_info { attlist_author_info, author+ }
attlist_author_info &= empty
```

[Child elements]

<author> Registers each author. Its syntax is given in Relax NG compact format below and explained in the following text.

```
author =
  element author {
    attlist_author,
    (personal_name | organization_name),
```

```
    address_info?,  
    author_other_info?  
  }  
  attlist_author &=  
  [ a:defaultValue = "author" ]  
  attribute role {  
    "author"  
    | "editor"  
    | "translator"  
    | "supervisor"  
    | "designer"  
    | "photographer"  
    | "illustrator"  
    | "binder"  
    | "planner"  
    | "other"  
  }?
```

[Attribute]

role: Defines the role of the person mentioned. The possible values are listed below. If omitted, it defaults to "author".

"author", "editor", "translator", "supervisor", "designer",
"photographer", "illustrator", "binder", "planner", "other"

[Child elements]

<personal_name> / <organization_name>

Records the author's name in one of these two elements, according to whether the author is an individual or an organization, respectively written as a Personal_name or Organization_name.

The details of the <personal_name> and <organization_name> elements are given in A.3.20 and A.3.21, respectively.

<address_info> Address of the author, written in the Address data format. May be omitted. The details of the <address_info> element are given in A.3.22.

<author_other_info> Other information related to the author may be stored as an External character string in this element. May be omitted. If the information is to be displayed, the space character (U+0020), and the linefeed and carriage return characters (U+000D, U+000A, U+000D+U+000A) are to be displayed as is, but the tabulation character (U+0009) is to be displayed as if it were a space. Its syntax is given in Relax NG compact format below.

```
author_other_info =
  element author_other_info { attlist_author_other_info, TextWithGaiji }
attlist_author_other_info &=
  [ a:defaultValue = "preserve" ]
attribute xml:space { "default" | "preserve" }?
```

<publisher_info> Stores the information related to the publisher. May be omitted. The <publisher_info> element has the following child elements. If <publisher_info> is specified, at least one of <publisher> or <publisher_office> shall be specified as well. Its syntax is given in Relax NG compact format below and explained in the following text.

```
publisher_info =
  element publisher_info {
    attlist_publisher_info,
    ((publisher, publisher_office) | publisher | publisher_office),
    publisher_other_info?
  }
attlist_publisher_info &= empty
```

[Child elements]

<publisher> Stores the information about the publisher if it is an individual. May be omitted. Its syntax is given in Relax NG compact format below and explained in the following text.

```
publisher =
  element publisher { attlist_publisher, publisher_name, address_info? }
attlist_publisher &= empty
```

[Child elements]

<publisher_name> The name of the publisher is recorded as an External character string in this element. Shall not be omitted. Its syntax is given in Relax NG compact format below and explained in the following text.

```
publisher_name =
  element publisher_name { attlist_publisher_name, TextWithGaiji }
attlist_publisher_name &= attribute reading { text }?
```

[Attribute]

reading: gives the pronunciation of the publisher's name as a Reading. May be omitted.

<address_info> Stores the address of the publisher, in the standard Address format. May be omitted. The details of the <address_info> element are given in A.3.22.

<publisher_office> Stores the information about the publisher if it is a company. May be omitted. Its syntax is given in Relax NG compact format below and explained in the following text.

publisher_office =

element publisher_office {

attlist_publisher_office, organization_name, address_info?

}

attlist_publisher_office &= attribute publisher_code { text }?

[Attribute]

publisher_code: Records the publisher's ID. May be omitted.

[Child elements]

<organization_name> The organization's name is recorded in the element. May be omitted. The details of the <organization_name> element are given in A.3.21.

<address_info> The publisher's address is recorded in this element, in the standard Address format. May be omitted. The details of the <address_info> element are given in A.3.22.

<publisher_other_info> Other information related to the publisher may be stored as an External character string in this element. May be omitted. If the information is to be displayed, the space character (U+0020), and the linefeed and carriage return characters (U+000D, U+000A, U+000D+U+000A) are to be displayed as is, but the tabulation character (U+0009) is to be displayed as if it were a space. Its syntax is given in Relax NG compact format below.

publisher_other_info =

element publisher_other_info {

attlist_publisher_other_info, TextWithGaiji

}

attlist_publisher_other_info &=

```
[ a:defaultValue = "preserve" ]
attribute xml:space { "default" | "preserve" }?
```

<seller_info> Stores the information related to the seller. May be omitted. If <seller_info> is specified, at least one of <seller> or <seller_office> shall be specified as well. Its syntax is given in Relax NG compact format below and explained in the following text.

```
seller_info =
  element seller_info {
    attlist_seller_info,
    ((seller, seller_office) | seller | seller_office),
    seller_other_info?
  }
attlist_seller_info &= empty
```

[Child elements]

<seller> Stores the information about the seller if it is an individual. May be omitted. Its syntax is given in Relax NG compact format below and explained in the following text.

```
seller = element seller { attlist_seller, seller_name, address_info? }
attlist_seller &= empty
```

[Child elements]

<seller_name> The name of the seller is recorded as an External character string in this element. Shall not be omitted. Its syntax is given in Relax NG compact format below and explained in the following text.

```
seller_name = element seller_name { attlist_seller_name,
  TextWithGaiji }
attlist_seller_name &= attribute reading { text }?
```

[Attribute]

reading: gives the pronunciation of the seller's name as a Reading. May be omitted.

<address_info> Stores the address of the seller, in the standard Address format. May be omitted. The details of the <address_info> element are given in A.3.22.

<seller_office> Stores the information about the seller if it is a company. May be omitted. Its syntax is given in Relax NG compact format below and explained in the following text.

```
seller_office =  
  element seller_office {  
    attlist_seller_office, organization_name, address_info?  
  }  
attlist_seller_office &= attribute seller_code { text }?
```

[Attribute]

seller_code: Records the seller's ID. May be omitted.

[Child elements]

<organization_name> Records the organization's name. May be omitted. The details of the <organization_name> element is given in A.3.21.

<address_info> Records the seller's address, written in the standard Address format. May be omitted. The details of the <address_info> element are given in A.3.22.

<seller_other_info> Other information related to the seller may be stored as an External character string in this element. May be omitted. If the information is to be displayed, the space character (U+0020), and the linefeed and carriage return characters (U+000D, U+000A, U+000D+U+000A) are to be displayed as is, but the tabulation character (U+0009) is to be displayed as if it were a space. Its syntax is given in Relax NG compact format below.

```
seller_other_info =  
  element seller_other_info { attlist_seller_other_info, TextWithGaiji }  
attlist_seller_other_info &=  
  [ a:defaultValue = "preserve" ]  
  attribute xml:space { "default" | "preserve" }?
```

<book_id_info> Records the book's identification number, such as its ISBN number. May be omitted. Its syntax is given in Relax NG compact format below and explained in the following text.

```
book_id_info = element book_id_info { attlist_book_id_info, book_id+ }  
attlist_book_id_info &= empty
```

[Child elements]

<book_id> Each type of identification number is stored in a <book_id> element, written as a Standard character string. When <book_id_info> is not omitted, there shall be at least one <book_id>. Its syntax is given in Relax NG compact format below and explained in the following text.

```
book_id = element book_id { attlist_book_id, text }  
attlist_book_id &= attribute type { text }
```

[Attribute]

type: Specifies the type of the identification number, such as "ISBN" for instance. Written as a Standard character string. Shall not be omitted.

<classification_info> Stores information on the classification of the book. May be omitted. Its syntax is given in Relax NG compact format below and explained in the following text.

```
classification_info =  
  element classification_info {  
    attlist_classification_info, classification+  
  }  
attlist_classification_info &= empty
```

[Child elements]

<classification> Each different type of classification is stored in a separate <classification> element. When <classification_info> is not omitted, there shall be at least one <classification>. It is stored as an External character string. Its syntax is given in Relax NG compact format below and explained in the following text.

```
classification =  
  element classification { attlist_classification, TextWithGaiji }  
attlist_classification &= attribute type { text }
```

[Attribute]

type: type of the classification used. Shall not be omitted.

<rating> Allows to rate the contents as violent, or adult. May be omitted. Its syntax is given in Relax NG compact format below and explained in the following text.

```
rating = element rating { attlist_rating, empty }  
  
attlist_rating &=
```

[a:defaultValue = "no"] attribute adult { Yes_No }?,

[a:defaultValue = "no"] attribute violence { Yes_No }?

[Attributes]

adult: Rates the contents as adult oriented materials. Possible values are "yes" or "no". Defaults to no in case of omission.

violence: Rates the contents as violent. Possible values are "yes" or "no". Defaults to no in case of omission.

<publication_place> The country of publication is recorded as a standard Country in this element. May be omitted.

<publication_date_info> Stores information regarding the publication date of the book. May be omitted. Its syntax is given in Relax NG compact format below and explained in the following text. Each relevant date is to be stored in separate instances of the following child element.

```
publication_date_info =
  element publication_date_info {
    attlist_publication_date_info, publication_date+
  }
```

attlist_publication_date_info &= empty

[Child elements]

<publication_date> Stores a date relevant to the publication, such as the publication date itself as well as other dates such as the printing date, or the beginning of sales date, etc. Each date is stored in the standard Date format. If <publication_date_info> is not omitted, there shall be at least one <publication_date>. Its syntax is given in Relax NG compact format below and explained in the following text.

```
publication_date =
  element publication_date { attlist_publication_date, text }
attlist_publication_date &=
  [ a:defaultValue = "publish" ] attribute type { "publish" | "sale" }?
```

[Attribute]

type: Specifies what type of date it is. "publish" if it is the publication date and "sale" if it is the beginning of sales date. If omitted, it will default to "publish".

<net_price_info> Defines the price of the book. May be omitted, if the price is open, or not set. Its syntax is given in Relax NG compact format below and explained in the following text. More than one price, classified by currency and country, may be stored in separate instances of the following child element.

```
net_price_info =
  element net_price_info { attlist_net_price_info, net_price+ }
attlist_net_price_info &= empty
```

[Child elements]

<net_price> Stores a price specific to one country and currency, written as a Standard character string. If <net_price_info> is not omitted, there shall be at least one <net_price>. Its syntax is given in Relax NG compact format below and explained in the following text.

```
net_price = element net_price { attlist_net_price, text }
```

```
attlist_net_price &=
```

```
  attribute country { text }?,
```

```
  attribute unit { text },
```

```
  attribute other_info { text }?
```

[Attributes]

country: Defines the country in which this price should apply, in the standard Country data format. If omitted, it applies to all countries.

unit: Defines the currency, as a Standard character string. Shall not be omitted.

other_info: Other information, written as a Standard character string. May be omitted.

<book_abstract> An abstract of the book, written as an External character string, is recorded in this element. May be omitted. If the information is to be displayed, the space character (U+0020), and the linefeed and carriage return characters (U+000D, U+000A, U+000D+U+000A) are to be displayed as is, but the tabulation character (U+0009) is to be displayed as if it were a space. Its syntax is given in Relax NG compact format below.

```
book_abstract =
```

```
  element book_abstract { attlist_book_abstract, TextWithGaiji }
```

```
attlist_book_abstract &=
```

```
  [ a:defaultValue = "preserve" ]
```

```
  attribute xml:space { "default" | "preserve" }?
```

<front_cover_image> Defines the image to use as front cover image, by recording its location as a standard Filename. Currently, jpeg and png images are supported. May be omitted. The file type as defined by it shall be checked against the type attribute before opening the file. Its syntax is given in Relax NG compact format below and explained in the following text.

```
front_cover_image =
```

```
  element front_cover_image { attlist_front_cover_image, text }
```

```
attlist_front_cover_image &= attribute type { text }
```

[Attribute]

type: Defines the type of the image, by giving its MIME type, for example "image/jpeg". Shall not be omitted.

<spine_cover_image> Define the image to use as a spine image. Follows the same rules as <front_cover_image>. May be omitted. Its syntax is given in Relax NG compact format below.

spine_cover_image =
 element spine_cover_image { attlist_spine_cover_image, text }
attlist_spine_cover_image &= attribute type { text }

<keyword_list> Records a list of keywords related to the book's data, written in the Keyword syntax. May be omitted. Its syntax is given in Relax NG compact format below.

keyword_list = element keyword_list { attlist_keyword_list, keyword+ }
attlist_keyword_list &= empty

<other_book_info> Other information related to the book may be stored as an External character string in this element. May be omitted. If the information is to be displayed, the space character (U+0020), and the linefeed and carriage return characters (U+000D, U+000A, U+000D+U+000A) are to be displayed as is, but the tabulation character (U+0009) is to be displayed as if it were a space. Its syntax is given in Relax NG compact format below.

other_book_info =
 element other_book_info { attlist_other_book_info, TextWithGaiji }
attlist_other_book_info &=
 [a:defaultValue = "preserve"]
 attribute xml:space { "default" | "preserve" }?

Example:

```
<book_info>
  <title_info>
    <series_title>Dummy books</series_title>
    <title>the dummy book of nonsense</title>
    <edition_info>2000/01/01 first edition,
                  2005/01/01 second edition</edition_info>
  </title_info>
  <author_info>
    <author role="author">
      <personal_name>
        <first_name>John</first_name>
        <last_name>Smith</last_name>
      </personal_name>
      <address_info>
        <mail_address>john.smith@abcd.com</mail_address>
        <website>http://www.abcd.com/~jsmith/</website>
      </address_info>
    </author>
  </author_info>
  <publisher_info>
    <publisher_office>
      <organization_name>abcd corporation</organization_name>
      <address_info>
        <postal_code>100-1000</postal_code>
        <address>1 main street, Foobar city, Japan</address>
      </address_info>
    </publisher_office>
  </publisher_info>
  <book_id_info>
    <book_id type="ISBN">xxx-x-xxxx-xxxx-x</book_id>
    <book_id type="Japanese_ID_number">454745-7</book_id>
  </book_id_info>
  <classification_info>
    <classification type="Japanese_C_CODE">2143</classification>
  </classification_info>
  <rating adult="no" violence="no"/>
  <publication_place>jpn</publication_place>
  <publication_date_info>
    <publication_date type="publish">2005</publication_date>
  </publication_date_info>
  <net_price_info>
    <net_price country="jpn" unit="yen">1200</net_price>
  </net_price_info>
  <book_abstract>This book doesn't talk about anything special.</book_abstract>
  <front_cover_image type="image/png">xxx.png</front_cover_image>
  <spine_cover_image type="image/png">yyy.png</spine_cover_image>
  <keyword_list>
    <keyword>dummy</keyword>
    <keyword>nonsense</keyword>
  </keyword_list>
</book_info>
```

A.4.3 Content management module <body_module>

A.4.3.1 General

The content management module (<body_module>) is in charge of coordinating the contents data into making the actual document. Its syntax is given in Relax NG compact format below and explained in the following text.

body_module =

element body_module { attlist_body_module, flow_type_body }

attlist_body_module &= empty

[Attribute]

None.

[Child elements]

<flow_type_body> Handles the flowing contents' data. See A.4.3.2 for details.

Example:

```
<body_module>
  <flow_type_body>
    <flow_entry>
      <flow_data ...> ... </flow_data>
      <flow_data ...> ... </flow_data>
      ...
    </flow_entry>
    <special_page_link>
      <special_page kind="contents">PG0001</special_page>
      <special_page kind="body">PG0002</special_page>
    </special_page_link>
    <search_table>
      ...
    </search_table>
  </flow_type_body>
</body_module>
```

A.4.3.2 Flowing content data <flow_type_body>

A.4.3.2.1 <flow_type_body> element

The <flow_type_body> element handles the flowing contents' data. Its syntax is given in Relax NG compact format below and explained in the following text.

flow_type_body =

element flow_type_body {

```
attlist_flow_type_body,

flow_entry,

special_page_link?,

search_table?

}

attlist_flow_type_body &= empty
```

[Attribute]

None.

[Child elements]

<flow_entry>	Registers the flow data to be used as the main text's flowing content. There shall only be one instance of this element. Shall not be omitted. See A.4.3.2.2 for details.
<special_page_link>	Special page data. Allows specifying the position in the flowing content of some often needed pages, such as the index, or the beginning of the main content, for easy reference. May be omitted. Written as described in A.4.3.2.5. Omission of this element means there is no special page.
<search_table>	Records the data needed to create a search table. May be omitted. Written as described in A.4.3.2.6. Omission of this element means there is no search table.

A.4.3.2.2 Flow data registering module <flow_entry>

The <flow_entry> element registers the flow data to be used as the main text's flowing content. Its syntax is given in Relax NG compact format below and explained in the following text.

```
flow_entry =

element flow_entry {

    attlist_flow_entry, flow_default_attribute?, flow_data+

}

attlist_flow_entry &= empty
```

[Child elements]

<flow_default_attribute>	Sets the default attributes to be used to display each flow data, as defined by the following <flow_data> element. May be omitted. See
--------------------------	--

A.4.3.2.3 for details. When omitted, the viewer should behave as if all of its attributes and child elements were set to their default value.

<flow_data>

Registers information on each flow data. There is a one to one relation between the number of <flow_data> elements and actual flow data to be recorded/displayed. There shall be one or more instances of this element. The order in which flow data are recorded determines the display order. See A.4.3.2.4 for details.

A.4.3.2.3 Flow data default attribute module <flow_default_attribute>

Sets the default values of attributes that will be used to display each flow data (see A.4.3.2.4), of which the main flowing content is composed. The values set in this element will serve as default values for all the flows which have a text object, a search screen object or dictionary data object as main data.

Part of the values that may be set in this element may also be set locally in the text object instance of each content data. In order to set a default value for particular content data, the <text_default_attribute> of the text object instances (see A.4.6.2) or <dict_default_attribute> of the dictionary data object designated by the content data should be used. If the value defined by the <flow_default_attribute> element for the whole content data and the value set in an individual content data's <text_default_attribute> or <dict_default_attribute> are in conflict, the latter has priority. When neither is set, the behavior is not defined by this annex, and depends on the viewer's default, or the user's preferences. Moreover, if the viewer does not implement the required method, or wishes to give priority to user settings, it may proceed without respecting the value defined in these default value elements, except when the following explanations state otherwise.

The syntax of the <flow_default_attribute> element is given in Relax NG compact format below and explained in the following text. If it is omitted, the viewer should behave as if all of its attributes and child elements were set to their own default value.

flow_default_attribute =

```
element flow_default_attribute {  
  
  attlist_flow_default_attribute,  
  
  flow_default_size?,  
  
  flow_default_font?,  
  
  flow_default_background?,  
  
  flow_default_line_breaking_method?  
  
}
```

attlist_flow_default_attribute &=

```
attribute baseline { BaseLine }?,  
  
attribute view_type { ViewType }?,  
  
[ a:defaultValue = "IPA" ] attribute phonetic_notation { text }?
```

[Attributes]

- baseline:** Defines the orientation of the baseline (and therefore of the text) for each flow. The possible values are listed below. May be omitted. If neither this attribute nor the corresponding option in each content data (the `baseline` attribute of the `<text_default_attribute>` element in text object instance or `<dict_default_attribute>` in dictionary data object instance, see A.4.6.2.2 and A.4.6.3.2 for details) is set, the default value depends on the viewer.
- `"right"`: The writing direction is horizontal (left to right). However, the direction can be changed at the user's option.
 - `"right_only"`: The writing direction is horizontal (left to right), and the user cannot change the setting. However, if it is not supported by the viewer, this setting is not necessarily applied.
 - `"down"`: The writing direction is vertical (top to bottom). However, the direction can be changed at the user's option.
 - `"down_only"`: The writing direction is vertical (top to bottom), and the user cannot change the setting. However, if it is not supported by the viewer, this setting is not necessarily applied.
- view_type:** Defines the default screen orientation for each flow. The possible values are listed below. May be omitted. When omitted, the default value depends on the viewer.
- `"portrait"`: Chooses portrait (taller than wide) mode. However, it can be set to another direction at the user's option.
 - `"portrait_only"`: Chooses portrait (taller than wide) mode, and the user cannot change the setting. However, if the viewer cannot handle this screen orientation, this is not necessarily applied.
 - `"landscape"`: Chooses landscape (wider than tall) mode. However, it can be set to another direction at the user's option.
 - `"landscape_only"`: Chooses landscape (wider than tall) mode, and the user cannot change the setting. However, if the viewer cannot handle this screen orientation, this is not necessarily applied.
- phonetic_notation:** Defines the default pronunciation notation of the content. Defaults to "IPA". This is overridden by the phonetic notation attribute of `<pronunciation>` and `<headword>` elements inside such elements.

Example:

```
<flow_default_attribute base_line="right" view_type="portrait">
```

[Child elements]

- <flow_default_size>** Defines the default letter spacing, line pitch and margin size for all the flows of the content data. May be omitted. If it is, the viewer should behave as if all of its attributes were set to their default value. Its syntax is given in Relax NG compact format below and explained in the following text.

flow_default_size =

```
element flow_default_size { attlist_flow_default_size, empty }  
attlist_flow_default_size &=  
  attribute letter_spacing { FiveSize }?,  
  attribute line_pitch { FiveSize }?,  
  attribute margin { "big" | "medium" | "small" }?
```

[Attributes]

letter_spacing: Default letter spacing. The following values are allowed. May be omitted. If it is, the size is unspecified and depends on the viewer.

"maximum"
"big"
"medium"
"small"
"minimum"

The actual sizes of these 5 possibilities are viewer-dependent, as it depends on the capabilities of the underlying device.

line_pitch: Default line pitch. The following values are allowed. May be omitted. If it is, the size is unspecified and depends on the viewer.

"maximum"
"big"
"medium"
"small"
"minimum"

The actual sizes of these 5 possibilities are viewer-dependent, as they depend on the capabilities of the underlying device.

margin: Default margin size. The following values are allowed. May be omitted. If it is, the size is unspecified and depends on the viewer.

"big"
"medium"
"small"

The actual sizes of these 3 possibilities are viewer-dependent, as they depend on the capabilities of the underlying device.

<flow_default_font> Defines the default font name, size, and properties for all the flows of the content data. May be omitted. If an individual flow of the content data also defines it (with the <text_default_font> element of the text object instance, as defined in A.4.6.2.2 and <dict_default_font> element of the dictionary data object instance as defined in A.4.6.3.2, the individual values take precedence. If both are omitted, the behavior should correspond to the default value of the attributes listed below. Its syntax is given in Relax NG compact format below and explained in the following text.

```
flow_default_font =  
  element flow_default_font { attlist_flow_default_font, empty }
```

```
attlist_flow_default_font &=
  attribute fontname { text }?,
  attribute fontsize { FiveSize }?,
  attribute bold_flag { Yes_No }?,
  attribute color_space { "RGB" }?,
  attribute opacity { "100" }?,
  attribute color { text }?,
  attribute ruby_flag { "yes" | "yes_only" | "no" | "no_only" }?,
  attribute italic { Yes_No }?,
  attribute oblique { Yes_No }?,
  attribute small_caps { Yes_No }?,
  attribute family {
    "monospace" | "san-serif" | "serif" | "cursive" | "fantasy"
  }?,
  attribute ul_type { "rightscore" | "leftscore" | "throughscore" }?,
  attribute em_type {
    "rightscore"
    | "leftscore"
    | "throughscore"
    | "kendot"
    | "bold"
    | "italic"
    | "bold_italic"
    | "reverse"
    | "shade"
  }?
```

[Attributes]

fontname:	Default font name. More than one font may be specified. In that case, each font name should be separated by a comma (U+002C). For instance: <pre>fontname="Aaa sans serif,Bbb gothic"</pre> The viewer should use the first listed font that is available. May be omitted. If both this attribute and the individual flow's attribute are omitted, the default value depends on the viewer.
fontsize:	Default font size. The following values are allowed. May be omitted. If it is, the size is unspecified and depends on the viewer. <pre> "maximum" "big" "medium" "small" "minimum" </pre>

The actual letter sizes corresponding to these 5 possibilities are viewer-dependent, as it depends on the capabilities of the underlying device.

bold_flag: Specifies if the content data should be displayed in bold or not. If omitted, the behavior depends on the viewer. The acceptable values are:

"yes": display as bold

"no": display normally

If set to "yes", all characters shall be displayed in bold style, except, in a text object instance or a dictionary data object instance, those within a element with its bold_flag attribute set to "no" (see A.4.6.2.3 and A.4.6.3.4).

color_space, color, opacity:

Defines the font color to be used for the content data. Written in the standard color data type. If omitted, the value depends on the viewer.

ruby_flag: Defines whether ruby in the content data is to be viewed or not. When omitted, behavior depends on the viewer. The text to be displayed when ruby is turned on is the one included in the <ruby> element of text object instances or dictionary data object instances. The following values may be used.

"yes": Ruby should be displayed, but can be turned off by the user.

"yes_only": Ruby shall be displayed, and cannot be turned off by the user. However, this does not apply to viewers not able to display ruby.

"no": It is recommended that ruby should not be displayed, but can still be turned on by the user.

"no_only": Ruby shall not be displayed, and cannot be turned on by the user. However, this does not apply if the viewer is incapable

italic: Decides if the content data should be displayed in italic or not. If omitted, the behavior depends on the viewer. The acceptable values are:

"yes": display in italic

"no": display normally

If set to "yes", all characters shall be displayed in italic style, except, in a text object instance or a dictionary data object instance, those within a element with its italic attribute set to "no" (see A.4.6.2.3 and A.4.6.3.4).

oblique: Decides if the content data should be displayed in oblique or not. If omitted, the behavior depends on the viewer. The acceptable values are

"yes": display in oblique;

"no": display normally.

If set to "yes", all characters shall be displayed in oblique style, except, in a text object instance or a dictionary data object instance, those within a element with its oblique attribute set to "no" (see A.4.6.2.3 and A.4.6.3.4).