

# INTERNATIONAL STANDARD

Multimedia systems and equipment – Multimedia e-publishing and e-books –  
Interchange format for e-dictionaries

(standards.iteh.ai)

IEC 62605:2016

<https://standards.iteh.ai/catalog/standards/sist/f1dfcf30-7d4b-4144-a6f2-9502f113e645/iec-62605-2016>



## THIS PUBLICATION IS COPYRIGHT PROTECTED

Copyright © 2016 IEC, Geneva, Switzerland

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 either IEC or IEC's member National Committee in the country of the requester. If you have any questions about IEC copyright or have an enquiry about obtaining additional rights to this publication, please contact the address below or your local IEC member National Committee for further information.

IEC Central Office  
3, rue de Varembe  
CH-1211 Geneva 20  
Switzerland

Tel.: +41 22 919 02 11  
Fax: +41 22 919 03 00  
[info@iec.ch](mailto:info@iec.ch)  
[www.iec.ch](http://www.iec.ch)

### About the IEC

The International Electrotechnical Commission (IEC) is the leading global organization that prepares and publishes International Standards for all electrical, electronic and related technologies.

### About IEC publications

The technical content of IEC publications is kept under constant review by the IEC. Please make sure that you have the latest edition, a corrigenda or an amendment might have been published.

#### IEC Catalogue - [webstore.iec.ch/catalogue](http://webstore.iec.ch/catalogue)

The stand-alone application for consulting the entire bibliographical information on IEC International Standards, Technical Specifications, Technical Reports and other documents. Available for PC, Mac OS, Android Tablets and iPad.

#### IEC publications search - [www.iec.ch/searchpub](http://www.iec.ch/searchpub)

The advanced search enables to find IEC publications by a variety of criteria (reference number, text, technical committee,...). It also gives information on projects, replaced and withdrawn publications.

#### IEC Just Published - [webstore.iec.ch/justpublished](http://webstore.iec.ch/justpublished)

Stay up to date on all new IEC publications. Just Published details all new publications released. Available online and also once a month by email.

#### Electropedia - [www.electropedia.org](http://www.electropedia.org)

The world's leading online dictionary of electronic and electrical terms containing 20 000 terms and definitions in English and French, with equivalent terms in 15 additional languages. Also known as the International Electrotechnical Vocabulary (IEV) online.

#### IEC Glossary - [std.iec.ch/glossary](http://std.iec.ch/glossary)

65 000 electrotechnical terminology entries in English and French extracted from the Terms and Definitions clause of IEC publications issued since 2002. Some entries have been collected from earlier publications of IEC TC 37, 77, 86 and CISPR.

#### IEC Customer Service Centre - [webstore.iec.ch/csc](http://webstore.iec.ch/csc)

If you wish to give us your feedback on this publication or need further assistance, please contact the Customer Service Centre: [csc@iec.ch](mailto:csc@iec.ch).

IEC 62601-2016  
INTERNATIONAL STANDARD PREVIEW  
(standards.itec.ai)  
https://standards.itec.ai/catalog/standards  
9502f113e645/iec-62601-2016

# INTERNATIONAL STANDARD

---

**Multimedia systems and equipment – Multimedia e-publishing and e-books –  
Interchange format for e-dictionaries**

**STANDARD PREVIEW**  
(standards.iteh.ai)  
<https://standards.iteh.ai/catalog/standards/sist/f1dfcf30-7d4b-4144-a6f2-9502f113e645/iec-62605-2016>

INTERNATIONAL  
ELECTROTECHNICAL  
COMMISSION

---

ICS 33.160.60; 35.240.20; 35.240.30

ISBN 978-2-8322-3102-9

**Warning! Make sure that you obtained this publication from an authorized distributor.**

## CONTENTS

FOREWORD.....	5
INTRODUCTION.....	7
1 Scope.....	8
2 Normative references.....	8
3 Terms and definitions .....	8
4 Position and requirements for interchange format for e-dictionaries .....	8
4.1 Interchange format for e-dictionaries in contents creation/distribution model.....	8
4.2 Requirements for interchange format for e-dictionaries.....	9
5 File formats .....	9
6 Semantics .....	10
Annex A (normative) XMDF-LeXML format .....	11
A.1 General.....	11
A.2 Overview of the format's structure .....	11
A.3 Elements and attributes.....	12
A.3.1 General .....	12
A.3.2 Page_ID .....	12
A.3.3 Object_ID .....	12
A.3.4 Char_ID .....	12
A.3.5 Reading .....	12
A.3.6 Filename.....	13
A.3.7 Standard character .....	13
A.3.8 Standard character string .....	13
A.3.9 Extended character .....	14
A.3.10 Extended character string.....	14
A.3.11 External character .....	14
A.3.12 External character string .....	16
A.3.13 External extended character string .....	16
A.3.14 Coordinates .....	16
A.3.15 Polygonal_region .....	17
A.3.16 Color .....	17
A.3.17 Date .....	18
A.3.18 Time .....	18
A.3.19 Country.....	18
A.3.20 Personal_name .....	18
A.3.21 Organization_name .....	20
A.3.22 Address .....	20
A.3.23 Permission.....	22
A.3.24 Keyword .....	24
A.3.25 Telephone_number .....	24
A.3.26 Mail_address .....	25
A.4 Description format details .....	25
A.4.1 General .....	25
A.4.2 Book information modules <bvf>.....	26
A.4.3 Content management module <body_module>.....	41
A.4.4 Event info module <event_info> .....	59

A.4.5	Parts data module <parts_module> .....	66
A.4.6	Object instances .....	72
A.5	Available color names .....	119
A.6	Localization .....	120
A.6.1	Possible additions .....	120
A.6.2	Standard characters .....	120
A.6.3	Characters usable for reading .....	121
A.6.4	Line breaking methods .....	121
A.6.5	Sorting rules for <search_table_def> .....	122
A.6.6	Additional attributes for <enable_key_type> .....	123
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 .....	124
Annex B (normative)	LeXML format .....	160
B.1	General .....	160
B.2	Elements for content structure .....	160
B.2.1	Parameter entity definition .....	160
B.2.2	Root elements .....	162
B.2.3	headword-related elements .....	166
B.2.4	Main-text related elements .....	169
B.2.5	Subheadword related elements .....	174
B.2.6	Other block elements .....	178
B.2.7	Media related elements .....	182
B.2.8	Other structural elements .....	184
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 .....	213
B.4	Specification of the LeXML format in the DTD syntax .....	220
Bibliography	.....	234
Figure 1	– Contents creation/distribution model .....	8
Figure 2	– Contents creation/distribution model (modified) .....	9
Figure 3	– Relationship between concepts .....	9
Figure A.1	– XML tree structure .....	11
Figure A.2	– Example of valign="middle" .....	74
Figure A.3	– Example of dropped capital .....	80
Figure A.4	– Left and right margin of a paragraph .....	80
Figure A.5	– Horizontal writing in vertical text .....	88
Figure A.6	– Ruby .....	89
Figure A.7	– Example of search page object instance rendering .....	115

Table A.1 – Base characters for reading .....	13
Table A.2 – Standard character set .....	13
Table A.3 – Usable characters for a telephone number .....	25
Table A.4 – Characters usable for email addresses .....	25
Table A.5 – Characters usable for the lookup key .....	58
Table A.6 – Color names.....	120
Table A.7 – Examples of additional standard character sets.....	121
Table A.8 – Example of additional characters usable for readings .....	121
Table A.9 – Example of additional sorting rules .....	122
Table A.10 – Example of additional language specific attributes for <enable_key_type> .....	123

## iTeh STANDARD PREVIEW (standards.iteh.ai)

[IEC 62605:2016](https://standards.iteh.ai/catalog/standards/sist/f1dfcf30-7d4b-4144-a6f2-9502f113e645/iec-62605-2016)

<https://standards.iteh.ai/catalog/standards/sist/f1dfcf30-7d4b-4144-a6f2-9502f113e645/iec-62605-2016>

## INTERNATIONAL ELECTROTECHNICAL COMMISSION

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

## FOREWORD

- 1) The International Electrotechnical Commission (IEC) is a worldwide organization for standardization comprising all national electrotechnical committees (IEC National Committees). The object of IEC is to promote international co-operation on all questions concerning standardization in the electrical and electronic fields. To this end and in addition to other activities, IEC publishes International Standards, Technical Specifications, Technical Reports, Publicly Available Specifications (PAS) and Guides (hereafter referred to as "IEC Publication(s)"). Their preparation is entrusted to technical committees; any IEC National Committee interested in the subject dealt with may participate in this preparatory work. International, governmental and non-governmental organizations liaising with the IEC also participate in this preparation. IEC collaborates closely with the International Organization for Standardization (ISO) in accordance with conditions determined by agreement between the two organizations.
- 2) The formal decisions or agreements of IEC on technical matters express, as nearly as possible, an international consensus of opinion on the relevant subjects since each technical committee has representation from all interested IEC National Committees.
- 3) IEC Publications have the form of recommendations for international use and are accepted by IEC National Committees in that sense. While all reasonable efforts are made to ensure that the technical content of IEC Publications is accurate, IEC cannot be held responsible for the way in which they are used or for any misinterpretation by any end user.
- 4) In order to promote international uniformity, IEC National Committees undertake to apply IEC Publications transparently to the maximum extent possible in their national and regional publications. Any divergence between any IEC Publication and the corresponding national or regional publication shall be clearly indicated in the latter.
- 5) IEC itself does not provide any attestation of conformity. Independent certification bodies provide conformity assessment services and, in some areas, access to IEC marks of conformity. IEC is not responsible for any services carried out by independent certification bodies.
- 6) All users should ensure that they have the latest edition of this publication.
- 7) No liability shall attach to IEC or its directors, employees, servants or agents including individual experts and members of its technical committees and IEC National Committees for any personal injury, property damage or other damage of any nature whatsoever, whether direct or indirect, or for costs (including legal fees) and expenses arising out of the publication, use of, or reliance upon, this IEC Publication or any other IEC Publications.
- 8) Attention is drawn to the Normative references cited in this publication. Use of the referenced publications is indispensable for the correct application of this publication.

International Standard IEC 62605 has been prepared by technical area 10: Multimedia e-publishing and e-book technologies, of IEC technical committee 100: Audio, video and multimedia systems and equipment.

This second edition cancels and replaces the first edition published in 2011. This edition constitutes a technical revision.

This edition includes the following significant technical changes with respect to the previous edition.

- a) Ref element is added to facilitate cross reference between entries.
- b) A new version of LeXML format, which is one of the base formats of the first edition, has been expanded and becomes Annex B. (The existing format becomes Annex A.)

The text of this standard is based on the following documents:

CDV	Report on voting
100/2430/CDV	100/2506/RVC

Full information on the voting for the approval of this standard can be found in the report on voting indicated in the above table.

This publication has been drafted in accordance with the ISO/IEC Directives, Part 2.

The committee has decided that the contents of this publication will remain unchanged until the stability date indicated on the IEC website under "<http://webstore.iec.ch>" in the data related to the specific publication. At this date, the publication will be

- reconfirmed,
- withdrawn,
- replaced by a revised edition, or
- amended.

A bilingual version of this publication may be issued at a later date.

## iTeh STANDARD PREVIEW (standards.iteh.ai)

[IEC 62605:2016](#)

<https://standards.iteh.ai/catalog/standards/sist/f1dfcf30-7d4b-4144-a6f2-9502f113e645/iec-62605-2016>



## 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

iTeh STANDARD PREVIEW

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.

[IEC 62605:2016](https://standards.iteh.ai/catalog/standards/sist/f1dfc30-7d4b-4144-a6f2-9502113c647/iec-62605-2016)

<https://standards.iteh.ai/catalog/standards/sist/f1dfc30-7d4b-4144-a6f2-9502113c647/iec-62605-2016>

ISO ([www.iso.org/patents](http://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

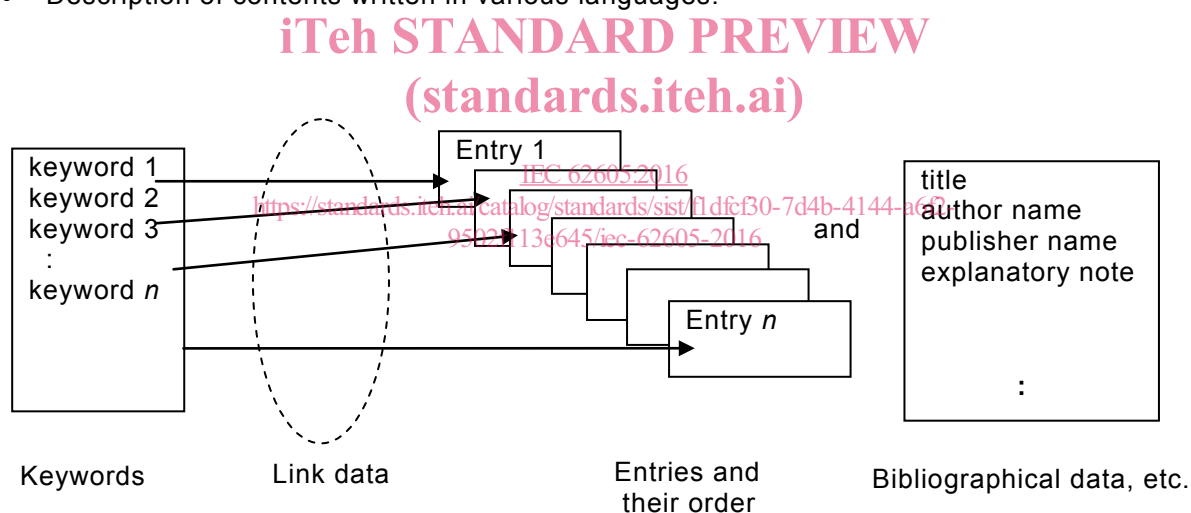
**Figure 2 – Contents creation/distribution model (modified)**

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.



IEC

**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.

**iTeh STANDARD PREVIEW**  
**(standards.iteh.ai)**

[IEC 62605:2016](https://standards.iteh.ai/catalog/standards/sist/f1dfcf30-7d4b-4144-a6f2-9502f113e645/iec-62605-2016)

<https://standards.iteh.ai/catalog/standards/sist/f1dfcf30-7d4b-4144-a6f2-9502f113e645/iec-62605-2016>

## Annex A (normative)

### X MDF-LeXML format

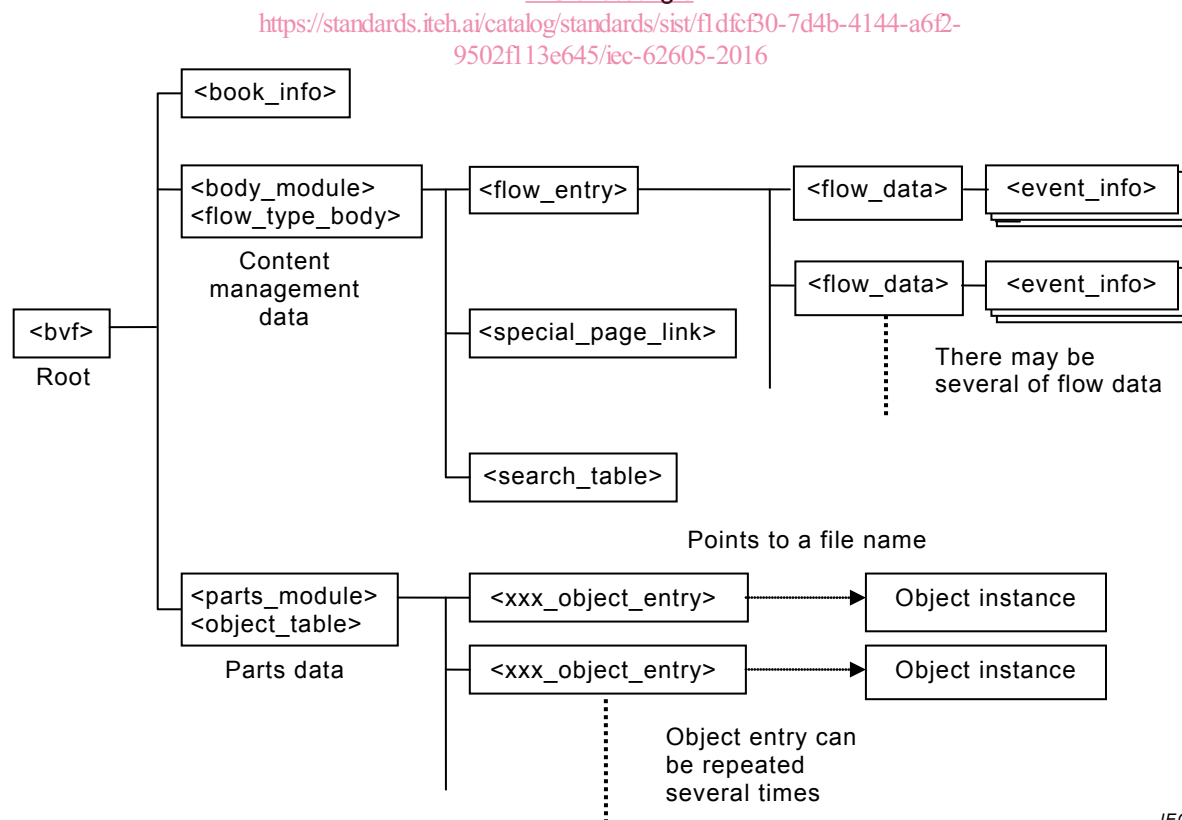
#### A.1 General

The XMDF-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 <sup>a</sup>
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)
<sup>a</sup> [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 ASCII 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:

IEC 62605:2016  
<https://standards.iteh.ai/catalog/standards/sist/f1df-f30-7d4b-4144-a6f2-657621c32c15/sist/62605-2016>

```
<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.