INTERNATIONAL STANDARD

ISO/IEC 29500-1

Third edition
2012-09-01

Information technology — Document description and processing languages — Office Open XML File Formats —

Part 1: Fundamentals and Markup Language Reference

Technologies de l'information — Description des documents et langages de traitement — Formats de fichier "Office Open XML" —

Partie 1: Principes essentiels et référence de langage de balisage

Reference number
ISO/IEC 29500-1:2012(E)

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Published in Switzerland

Table of Contents

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work. In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC 1.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of the joint technical committee is to prepare International Standards. Draft International Standards adopted by the joint technical committee are circulated to national bodies for voting. Publication as an International Standard requires approval by at least 75 % of the national bodies casting a vote.

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

ISO/IEC 29500-1 was prepared by ISO/IEC JTC 1, Information technology, Subcommittee SC 34, Document description and processing languages.

This third edition cancels and replaces the second edition (ISO/IEC 29500-1:2011), which has been technically revised by incorporation of the Amendment ISO/IEC 29500-1:2011/Amd.1:2012 and the Technical Corrigendum ISO/IEC 29500-1:2011/Cor.1:2012.

ISO/IEC 29500 consists of the following parts, under the general title *Information technology — Document description and processing languages — Office Open XML File Formats*:

- *Part 1: Fundamentals and Markup Language Reference*
- *Part 2: Open Packaging Conventions*
- *Part 3: Markup Compatibility and Extensibility*
- *Part 4: Transitional Migration Features*

Annexes A, G and H form a normative part of this Part of ISO/IEC 29500. Annexes B–F and I–N are for information only.

This Part of ISO/IEC 29500 includes five annexes (Annex A, Annex B, Annex F, Annex G, and Annex H) that refer to data files provided in electronic form.

The document representation formats defined by this Part are different from the formats defined in the corresponding Part of ECMA-376:2006. Some of the differences are reflected in schema changes, as shown in Annex M of this Part.

Introduction

ISO/IEC 29500 specifies a family of XML schemas, collectively called *Office Open XML*, which define the XML vocabularies for word-processing, spreadsheet, and presentation documents, as well as the packaging of documents that conform to these schemas.

The goal is to enable the implementation of the Office Open XML formats by the widest set of tools and platforms, fostering interoperability across office productivity applications and line-of-business systems, as well as to support and strengthen document archival and preservation, all in a way that is fully compatible with the existing corpus of Microsoft Office documents.

The following organizations have participated in the creation of ISO/IEC 29500 and their contributions are gratefully acknowledged:

Apple, Barclays Capital, BP, The British Library, Essilor, Intel, Microsoft, NextPage, Novell, Statoil, Toshiba, and the United States Library of Congress

Information technology — Document description and processing languages — Office Open XML File Formats

Part 1:

Fundamentals and Markup Language Reference

1. Scope

ISO/IEC 29500 defines a set of XML vocabularies for representing word-processing documents, spreadsheets and presentations. On the one hand, the goal of ISO/IEC 29500 is to be capable of faithfully representing the pre-existing corpus of word-processing documents, spreadsheets and presentations that had been produced by the Microsoft Office applications (from Microsoft Office 97 to Microsoft Office 2008, inclusive) at the date of the creation of ISO/IEC 29500. It also specifies requirements for Office Open XML consumers and producers. On the other hand, the goal is to facilitate extensibility and interoperability by enabling implementations by multiple vendors and on multiple platforms.

This Part of ISO/IEC29500 specifies concepts for documents and applications of both strict and transitional conformance.

2. Conformance

2.1 Document Conformance

A document of conformance class Office Open XML Strict shall be a package of conformance class OPC, as specified in ISO/IEC 29500-2, for which all the following shall hold:

- The document obeys all constraints specified in this Part of ISO/IEC 29500
- The document is of category Wordprocessing, Spreadsheet, or Presentation, as defined in §4
- For each OPC Part of the document of the types listed in §11.3, §12.3, §13.3, §14.2 or §15.2, all the following shall hold:
    - i. The part is of conformance class MCE, as specified in ISO/IEC 29500-3
    - ii. After the removal of any extensions using the mechanisms in ISO/IEC 29500-3, the part is valid against the strict W3C XML Schema (Appendix A)

This Part of ISO/IEC 29500 uses the following further terms to refer to documents of conformance class Office Open XML Strict:

- *WML Strict*, if the document is of category Wordprocessing
- *SML Strict*, if the document is of category Spreadsheet
- *PML Strict*, if the document is of category Presentation

2.2 Application Conformance

Application conformance incorporates both syntax and semantics.

- A conforming consumer shall not reject any conforming documents of at least one document conformance class.
- A conforming producer shall be able to produce conforming documents of at least one document conformance class.
- A conforming application shall treat the information in Office Open XML documents in a manner consistent with the semantic definitions given in ISO/IEC 29500. An application's intended behavior need not require that application to process all of the information in an Office Open XML document. However, the information that it does process shall be processed in a manner that is consistent with the semantic definitions given in ISO/IEC 29500.

[*Note*: This note illustrates the third bullet above. Conforming applications might serve various functions. Examples include a viewer, an editor, and a back-end processor. Here is an illustration of how the third bullet applies to each of those examples:

- If a conforming viewer supports a given feature, then when it displays information using that feature, it respects the semantics of that feature as described in the Standard.

- If a conforming editor supports a given feature, then when it provides its user with an interface for manipulating information using that feature, it respects the semantics of that feature as described in the Standard.
- If a conforming back-end processor supports a given feature, then when that processor transforms or assembles information involving that feature, that processor respects the semantics of that feature as described in the Standard.

end note]

This Part of ISO/IEC 29500 defines the following application conformance classes:

- *WML Strict*, if the application is a conforming application that is a consumer or producer of documents having conformance class WML Strict.
- *SML Strict*, if the application is a conforming application that is a consumer or producer of documents having conformance class SML Strict.
- *PML Strict*, if the application is a conforming application that is a consumer or producer of documents having conformance class PML Strict.

Conformance can also involve the use of application descriptions; see §2.3 for details.

2.3 Application Descriptions

An application can be defined as conforming to zero or more *application descriptions* in a particular conformance class.

The application descriptions defined within ISO/IEC 29500 are:

- Base
- Full

[*Note*: These application descriptions should not be taken as limiting the ability of an application provider to create innovative applications. They are intended as a mechanism for labelling applications rather than for restricting their capabilities. The intention is to promote interoperability between different applications that share the same conformance class. Application descriptions are orthogonal to the conformance of the documents produced by those applications. For example, a tool used for automated translation of documents might have an application description of “Base” but will still produce fully conformant documents. *end note*]

The application descriptions are determined in terms of an application’s semantic understanding of particular features. *Semantic understanding* is to be interpreted in that an application shall treat the information in Office Open XML documents in a manner consistent with the semantic definitions given in ISO/IEC 29500.

Each application description is identified by a URI.

The application descriptions are defined in the following subclauses.

2.3.1 Base Application Description

Description URI: http://purl.oclc.org/ooxml/descriptions/base

An application conforming to this description has a semantic understanding of at least one feature within its conformance class.

[*Note*: In addition, applications that include a user interface are strongly recommended to support all accessibility features appropriate to that user interface. *end note*]

2.3.2 Full Application Description

Description URI: http://purl.oclc.org/ooxml/descriptions/full

An application conforming to this description has a semantic understanding of every feature within its conformance class.

2.3.3 Additional Application Descriptions

It is expected that additional application descriptions will be defined within the maintenance process for ISO/IEC 29500. It is also expected that third parties might define their own application descriptions; for example to inform their procurement decisions, or to deal with domains such as accessibility.

[*Note*: A possible application description would be a "standard" application description for a wordprocessing application. This could be created by taking the intersection of the features available in common wordprocessing applications such as Word 2000, OpenOffice 2, WordPerfect, and iWork Pages. In addition, it could define formats such as specific image and video formats required to be supported to conform to the description. Similar descriptions could be created for spreadsheet applications and presentation applications. Such a description would promote interoperability between applications implementing OOXML. It would also promote interoperability between applications implementing OOXML and applications implementing other document formats such as ISO/IEC 26300. *end note*]

Application descriptions are not required to be strict subsets of each other. An application can simultaneously conform to multiple application descriptions.

Any such newly created description shall enumerate the features that are required for conformance to it. Such a description should provide a machine-processable schema, preferably using a standard such as ISO/IEC 19757.

[*Note*: If the application conforming to a description is a document consumer, it should be able to consume any document that respects such a schema associated with the description. If the application is a document producer, any document produced by that application should respect the schema of the description. *end note*]

Any such description should be identified using a URI, in a similar manner to the names used for application descriptions within ISO/IEC 29500.

[*Note*: For the convenience of users of the description, it is recommended that creators of a description should make a human- or machine-readable form of that description available at a URL corresponding to the description URI. *end note*]

2.3.4 Representation of Application Descriptions within Documents

An application description is related to applications, rather than to document conformance. Therefore, there is no normative mechanism for representing an application description within a document.

[*Note*: It is recommended that implementers wishing to represent an application description within a document use the standard metadata mechanism for Office Open XML. *end note*]

2.4 Interoperability Guidelines

[*Guidance*: The following interoperability guidelines incorporate semantics]

For the guidelines to be meaningful, a software application should be accompanied by documentation that describes what subset of ISO/IEC 29500 it supports. The documentation should highlight any behaviors that would, without that documentation, appear to violate the semantics of document XML elements. Together, the application and documentation should satisfy the following conditions.

1. The application need not implement operations on all XML elements defined in ISO/IEC 29500. However, if it does implement an operation on a given XML element, then that operation should use semantics for that XML element that are consistent with ISO/IEC 29500.
2. If the application moves, adds, modifies, or removes XML element instances with the effect of altering document semantics, it should declare the behavior in its documentation.

The following scenarios illustrate these guidelines.

- A presentation editor that interprets the preset shape geometry "`rect`" as an ellipse does not observe the first guideline because it implements "`rect`" but with incorrect semantics.
- A batch spreadsheet processor that saves only computed values even if the originally consumed cells contain formulas, might satisfy the first condition, but does not observe the second because the editability of the formulas is part of the cells' semantics. To observe the second guideline, its documentation should describe the behavior.
- A batch tool that reads a word-processing document and reverses the order of text characters in every paragraph with "Title" style before saving it can be conforming even though ISO/IEC 29500 does not recommend this behavior. This tool's behavior would be to transform the title "Office Open XML" into "LMX nepO eciffO". Its documentation should declare its effect on such paragraphs.

The normative requirements in §2.1 imply that a conforming producer shall not write unescaped non-XML characters. As an implementation guideline, a conforming producer additionally should not write escaped non-XML characters. Doing so damages interoperability with existing XML-based standards such as SOAP and RDF.

