
**Document description and processing
languages — Office Open XML file
formats —**

Part 2:
Open packaging conventions

*Description des documents et langages de traitement — Formats de
fichier "Office Open XML" —*

Partie 2: Conventions de paquetage ouvert





COPYRIGHT PROTECTED DOCUMENT

© ISO/IEC 2021

All rights reserved. Unless otherwise specified, or required in the context of its implementation, no part of this publication may be reproduced or utilized otherwise in any form or by any means, electronic or mechanical, including photocopying, or posting on the internet or an intranet, without prior written permission. Permission can be requested from either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier; Geneva
Phone: +41 22 749 01 11
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Contents

	Page
Foreword	v
Introduction	vii
1 Scope	1
2 Normative references	1
3 Terms and definitions	2
3.1 Basics	2
3.2 Abstract package model	3
3.3 Physical package model	4
3.4 Digital signature and thumbnail	5
3.5 Implementations	5
3.6 Core properties	5
4 Conformance	5
5 Overview	5
6 Abstract package model	6
6.1 General	6
6.2 Parts	6
6.2.1 General	6
6.2.2 Part names	6
6.2.3 Media types	8
6.2.4 Growth hint	8
6.2.5 XML usage	8
6.3 Part addressing	8
6.3.1 General	8
6.3.2 Pack scheme	9
6.3.3 Resolving a pack IRI to a resource	10
6.3.4 Composing a pack IRI	10
6.3.5 Equivalence	11
6.4 Resolving relative references	11
6.4.1 General	11
6.4.2 Base IRIs	11
6.4.3 Examples	12
6.5 Relationships	14
6.5.1 General	14
6.5.2 Relationships part	14
6.5.3 Relationship markup	15
6.5.4 Examples	17
7 Physical package model	19
7.1 General	19
7.2 Physical mapping guidelines	19
7.2.1 Using features of physical formats	19
7.2.2 Mapped components	20
7.2.3 Mapping media types to parts	20
7.2.4 Interleaving	23
7.2.5 Mapping part names to physical package item names	23
7.3 Mapping to a ZIP file	25
7.3.1 General	25
7.3.2 Mapping part data	25
7.3.3 ZIP item names	26
7.3.4 Mapping logical item names to ZIP item names	26
7.3.5 Mapping ZIP item names to logical item names	26
7.3.6 ZIP package limitations	26
7.3.7 Mapping the Media Types stream	27

7.3.8	Mapping the growth hint.....	27
8	Core properties	27
8.1	General.....	27
8.2	Core Properties part.....	28
8.3	Core properties markup.....	28
8.3.1	General.....	28
8.3.2	Support for versioning and extensibility.....	29
8.3.3	<code>coreProperties</code> element.....	29
8.3.4	Core property elements.....	29
9	Thumbnails	32
10	Digital signatures	32
10.1	General.....	32
10.2	Overview of OPC-specific restrictions and extensions to “XML-Signature Syntax and Processing”.....	32
10.3	Choosing content to sign.....	32
10.4	Digital signature parts.....	33
10.4.1	General.....	33
10.4.2	Digital Signature Origin part.....	33
10.4.3	Digital Signature XML Signature part.....	33
10.4.4	Digital Signature Certificate part.....	34
10.5	Digital signature markup.....	34
10.5.1	General.....	34
10.5.2	Support for versioning and extensibility.....	34
10.5.3	<code>Signature</code> element.....	34
10.5.4	<code>SignedInfo</code> element.....	35
10.5.5	<code>CanonicalizationMethod</code> element.....	35
10.5.6	<code>SignatureMethod</code> element.....	35
10.5.7	<code>Reference</code> element.....	35
10.5.8	<code>Transform</code> element.....	36
10.5.9	<code>RelationshipReference</code> element.....	37
10.5.10	<code>RelationshipsGroupReference</code> element.....	37
10.5.11	<code>DigestMethod</code> element.....	37
10.5.12	<code>Object</code> element.....	38
10.5.13	<code>Manifest</code> element.....	38
10.5.14	<code>SignatureProperty</code> element.....	38
10.5.15	<code>SignatureTime</code> element.....	38
10.5.16	<code>Format</code> element.....	38
10.5.17	<code>Value</code> element.....	39
10.5.18	<code>XPath</code> element.....	39
10.6	Relationships transform algorithm.....	39
10.7	Digital signature example.....	40
10.8	Generating signatures.....	41
10.9	Validating signatures.....	43
	Annex A (informative) Preprocessing for generating relative references	45
	Annex B (normative) Constraints and clarifications on the use of ZIP features	46
	Annex C (normative) Schemas - W3C XML	55
	Annex D (informative) Schemas - RELAX NG	56
	Annex E (normative) Standard namespaces and media types	57
	Annex F (informative) Physical package model design considerations	58
	Annex G (informative) Differences between ISO/IEC 29500-2 and ECMA-376:2006	62
	Annex H (informative) Package example	63
	Bibliography	65

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.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of document should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives or www.iec.ch/members_experts/refdocs).

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. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents) or the IEC list of patent declarations received (see patents.iec.ch).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see www.iso.org/iso/foreword.html. In the IEC, see www.iec.ch/understanding-standards.

This document was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 34, *Document description and processing languages*.

This fourth edition cancels and replaces the third edition (ISO/IEC 29500-2:2012), which has been technically revised.

The main changes compared to the previous edition are as follows:

- Where appropriate, normative references have been updated to use undated or more recent versions of other standards.
- [Clause 3](#) (Terms and definitions) has been revised by removing terms not used by any normative clauses and then reorganizing terms into groups.
- The subclause for diagram notes (5.1 in the preceding editions) has been removed, since core properties are now defined by prose and schemas rather than by diagrams.
- The clause for acronyms and abbreviations (Clause 6 in the preceding editions) has been removed, since it does not make sense for an ISO/IEC standard to define "ISO" and "IEC".
- [Clause 6](#) (Abstract package model, Clause 8 in the previous edition) has been completely rewritten. In particular, (1) pack IRIs have been defined in this clause rather than in an annex, (2) a new subclause, "Resolving relative references", has been added; (3) part Relationships parts and package Relationships parts have been distinguished; (4) base IRIs have been clearly defined; and (5) handling of non-ASCII characters in part names has been clarified on the basis of RFC 3987.
- The option for media type to be an empty string has been removed, as this conflicts with the definition of media type in RFC 2046 and the existing regular expression defined in the schema referenced by [C.2](#).

ISO/IEC 29500-2:2021(E)

- [Clause 7](#) (Physical package model, Clause 9 in the previous edition) has been slightly revised. Interleaving has been introduced before logical item names. Percent-encoding and un-percent encoding of non-ASCII characters have been explicitly introduced in [7.3.4](#) and [7.3.5](#).
- [Clause 8](#) (Core properties, Clause 10 in the previous edition) has been rewritten by using prose and schemas rather than diagrams.
- [Clause 10](#) (Digital signatures, Clause 12 in the previous edition) has been thoroughly revised. In particular, this clause now makes clear a convention for the choice of algorithms for signature and digest methods, which reflects the ongoing development of algorithms since the first edition of this document.
- [Annex A](#) has been made informative.
- The normative annex that defined pack IRIs (Annex B in the preceding editions) has been dropped. Pack IRIs are now defined in [Clause 6](#).
- [Annex C](#) and [Annex D](#) (Annexes D and E in the preceding editions) no longer define schemas but reference externally defined schemas.
- Guidelines for meeting conformance requirements (Annex H in the preceding editions) have been dropped.
- Requirements around streaming consumption have been dropped.
- Wherever possible, requirements on programs have been rewritten as those on data.
- [Annex H](#) has been added to depict an example package.
- The Index (Annex J in the preceding editions) has been deleted.
- Bibliography has been added.

A list of all parts in the ISO/IEC 29500 series can be found on the ISO and IEC websites.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html and www.iec.ch/national-committees.

Introduction

ISO/IEC 29500 (all Parts) 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® Office¹⁾ documents.

This document includes two annexes ([Annex C](#) and [Annex D](#)) that refer to data files provided in electronic form.

The document representation formats defined by this document 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 G](#).

This fourth edition preserves all previous functionality and adds no new functionality.

1) This information is given for the convenience of users of this document and does not constitute an endorsement by ISO or IEC of the product named. Equivalent products may be used if they can be shown to lead to the same results.

Document description and processing languages — Office Open XML file formats —

Part 2: Open packaging conventions

1 Scope

This document defines a set of conventions for packaging one or more interrelated byte streams (parts) as a single resource (package). These conventions are applicable not only to Office Open XML specifications as described in ISO/IEC 29500-1 and ISO/IEC 29500-4, but also to other markup specifications.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ANSI/INCITS 4-1986 [R2017] - *Information Systems - Coded Character Sets - 7-Bit American National Standard Code For Information Interchange (7-Bit ASCII)*, American National Standards Institute (ANSI), 2017

FIPS 186-4, *Digital Signature Standard (DSS)*, National Institute of Standards and Technology, US Department of Commerce, July 2013

ISO/IEC 29500-3, *Information technology — Document description and processing languages — Office Open XML File Formats — Part 3: Markup Compatibility and Extensibility*

ISO/IEC 9594-8/ITU-T Rec. X.509, *Information technology — Open systems interconnection — Part 8: The Directory: Public-key and attribute certificate frameworks*

ISO 15836-1, *Information and documentation — The Dublin Core metadata element set — Part 1: Core elements*

ISO 15836-2, *Information and documentation — The Dublin Core metadata element set — Part 2: DCMI Properties and classes*

RFC 2046, *Multipurpose Internet Mail Extensions (MIME) Part Two: Media Types*, The Internet Society, November 1996, N. Freed and N. Borenstein. Available at <https://www.rfc-editor.org/info/rfc2046>

RFC 3986, *Uniform Resource Identifier (URI): Generic Syntax*, The Internet Society, January 2005, Berners-Lee, T., R. Fielding, and L. Masinter. Available at <https://www.rfc-editor.org/info/rfc3986>

RFC 3987, *Internationalized Resource Identifiers (IRIs)*, The Internet Society, January 2005, Duerst, M. and M. Suignard. Available at <https://www.rfc-editor.org/info/rfc3987>

RFC 5234, *Augmented BNF for Syntax Specifications: ABNF*, The Internet Society, January 2008, D. Crocker and P. Overell, (editors). Available at <https://www.rfc-editor.org/info/rfc5234>

RFC 6931, *Additional XML Security Uniform Resource Identifiers (URIs)*, The Internet Society, April 2013, D. Eastlake 3rd. Available at <https://www.rfc-editor.org/info/rfc6931>

ISO/IEC 29500-2:2021(E)

RFC 7231, *Hypertext Transfer Protocol (HTTP/1.1): Semantics and Content*, The Internet Society, June 2014, R. Fielding and J. Reschke. Available at <https://www.rfc-editor.org/info/rfc7231>

Unicode, *The Unicode Standard*, The Unicode Consortium. Available at <http://www.unicode.org/standard/standard.html>

The XML 1.0 specification, *Extensible Markup Language (XML) 1.0, Fourth Edition*. World Wide Web Consortium, 2006, Tim Bray, Jean Paoli, Eve Maler, C. M. Sperberg-McQueen, and François Yergeau (editors). Available at <http://www.w3.org/TR/2006/REC-xml-20060816/>²⁾

XML Namespaces, *Namespaces in XML 1.0 (Third Edition)*, 8 December 2009. World Wide Web Consortium, Tim Bray, Dave Hollander, Andrew Layman, and Richard Tobin (editors). Available at <http://www.w3.org/TR/2009/REC-xml-names-20091208/>

XML Base, *XML Base (Second Edition)*, World Wide Web Consortium, 28 January 2009. Jonathan Marsh and Richard Tobin (editors). Available at <https://www.w3.org/TR/2009/REC-xmlbase-20090128/>

W3C XML Schema Structures, *XML Schema Part 1: Structures (Second Edition)*, World Wide Web Consortium, 28 October 2004, Henry Thompson, David Beech, Murray Maloney and Noah Mendelsohn (editors). Available at <https://www.w3.org/TR/xmlschema-1/>

W3C XML Schema Datatypes, *XML Schema Part 2: Datatypes (Second Edition)*, World Wide Web Consortium, 28 October 2004, Paul Biron and Ashok Malhotra (editors). Available at <https://www.w3.org/TR/xmlschema-2/>

XML-Signature Syntax and Processing, World Wide Web Consortium, 12 February 2002, Donald Eastlake, Joseph Reagle and David Solo (editors). Available at <http://www.w3.org/TR/2002/REC-xmlsig-core-20020212/>

ZIP Appnote, *ZIP File Format Specification* Version 6.2.0, PKWARE Inc., 2004. Available at http://www.pkware.com/documents/APPNOTE/APPNOTE_6.2.0.txt

2) A further correction of the normative reference to XML to refer to the 5th Edition will be necessary when the related Reference Specifications to which this document also makes normative reference, and which also depend upon XML, such as XML Namespaces and XML Base, are all aligned with the 5th Edition.