INTERNATIONAL STANDARD

ISO/IEC 19757-11

First edition
2011-11-01

Information technology — Document Schema Definition Languages (DSDL) —

Part 11: Schema association

Technologies de l'information — Langages de définition de schéma de documents (DSDL) —
Partie 11: Association de schémas

Reference number
ISO/IEC 19757-11:2011(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

Contents

Foreword iv
Introduction v
1 Scope 1
2 Normative references 1
3 Conformance requirements 1
4 The xml-model processing instruction 2
Annex A (informative) **Examples** 4
Annex B (informative) **Suggested use of schematypens for determining schema language** 5

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 19757-11 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 34, *Document description and processing languages*.

ISO/IEC 19757 consists of the following parts, under the general title *Information technology — Document Schema Definition Languages (DSDL)*:

— *Part 1: Overview* [Technical Report]

— *Part 2: Regular-grammar-based validation — RELAX NG*

— *Part 3: Rule-based validation — Schematron*

— *Part 4: Namespace-based Validation Dispatching Language (NVDL)*

— *Part 5: Extensible Datatypes*

— *Part 7: Character Repertoire Description Language (CREPDL)*

— *Part 8: Document Semantics Renaming Language (DSRL)*

— *Part 9: Namespace and datatype declaration in Document Type Definitions (DTDs)*

— *Part 11: Schema association*

Introduction

There are several document schema definition languages in common use today that can be used to specify one or more validation processes performed against Extensible Markup Language (XML) documents. Some schema languages provide their own syntax for associating schemas with documents (DTD, W3C XML Schema) and some languages (RELAX NG, Schematron) do not provide schema association mechanisms at all. The purpose of this part of ISO/IEC 19757 is to define a common, schema-agnostic syntax for associating schema documents written in any schema definition language with a given XML document.

This part of ISO/IEC 19757 defines the syntax and processing expectations for an xml-model processing instruction. Such processing instructions associate one or more schemas with the XML document in which they are present. The associated schemas may be written in any schema definition language. Applications can use the associated schemas for any purpose including those such as document validation, content completion in interactive editors, or creating models for data binding. Presence of an xml-model processing instruction is not in itself an instruction to any processor to validate the document, nor is it a statement that the document is not to be processed without validation. It is a declarative statement of a relationship between the document and one or more external schemas. This part of ISO/IEC 19757 does not prescribe what, if anything, an application does with an xml-model processing instruction. The presence of an xml-model processing instruction referencing a Document Type Definition (DTD) does not affect the validity of the document which contains it.

This part of ISO/IEC 19757 is not meant as a replacement for other technologies that provide more general and indirect schema association features like Namespace-based Validation Dispatching Language (NVDL) and XProc. This part of ISO/IEC 19757 is complementary technology which can be used when it is necessary to store ad hoc schema associations directly inside an XML document.

The technical content of this part of ISO/IEC 19757 has been developed jointly with W3C. It is expected to be identical to that of W3C's *Associating Schemas with XML documents 1.0*.

Information technology — Document Schema Definition Languages (DSDL) —

Part 11: Schema association

1 Scope

This part of ISO/IEC 19757 allows schemas using any schema definition language to be associated with an XML document by including one or more processing instructions with a target of xml-model in the document's prolog.

2 Normative references

The following referenced documents are indispensable for the application 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.

[ASSOCSS] *Associating Style Sheets with XML documents 1.0 (Second Edition)*. W3C, 28 October 2010. Available at <http://www.w3.org/TR/2010/REC-xml-stylesheet-20101028/>

[IANACHARSET] *Character Sets*, IANA, May 2007. Available at <http://www.iana.org/assignments/character-sets>

[ISO/IEC 19757-3] ISO/IEC 19757-3:2006, *Information technology — Document Schema Definition Languages (DSDL) — Part 3: Rule-based validation — Schematron*

[IETF RFC3987] *Internationalized Resource Identifiers (IRIs)*, M. Dürst, M. Suignard. IETF, January 2005. Available at <http://www.ietf.org/rfc/rfc3987.txt>

[XML] *Extensible Markup Language (XML) 1.0 (Fifth Edition)*, T. Bray, J. Paoli, C. Sperberg-McQueen, E. Maler, F. Yergeau. W3C, November 2008. Available at <http://www.w3.org/TR/xml/>

3 Conformance requirements

All diagrams, examples, and notes in this part of ISO/IEC 19757 are non-normative, as are all sections explicitly marked non-normative. Everything else in this part of ISO/IEC 19757 is normative.

Documents

A document is considered to conform to this part of ISO/IEC 19757 if it satisfies all criteria in this part of ISO/IEC 19757 that apply to documents.

xml-model processors

XML defines an *application* as a software module which receives the information content of an XML document from an *XML processor*. An *xml-model processor* is such an application which processes XML *processing instructions* in accordance with this part of ISO/IEC 19757.

An xml-model processor may be part of a larger XML application, or may function independently. In either case, an *application* is the consumer of the pseudo-attribute information defined in this part of ISO/IEC 19757.

An xml-model processor is considered to be a conforming xml-model processor if it satisfies all criteria in this part of ISO/IEC 19757 that apply to xml-model processors; xml-model processors do not have to check or enforce any of the constraints on documents.

This part of ISO/IEC 19757 is defined with reference to the vocabulary for XML provided by the XML Information Set as well as the *rules for parsing pseudo-attributes from a string* as defined in the Associating Style Sheets with XML documents Recommendation *[ASSOCSS]*.

The productions in this part of ISO/IEC 19757 use the same notation as used in the XML Recommendation. Tokens in the grammar and terms used in this part of ISO/IEC 19757 that are not defined in this part of ISO/IEC 19757 are defined in the XML Recommendation *[XML]* or the Associating Style Sheets *[ASSOCSS]* Recommendation.

4 The xml-model processing instruction

A processing instruction information item is said to be a *potential xml-model processing instruction* if it has the [target] property xml-model and it is in the [children] property of a document information item and appears before the element information item of the document information item's [children] property.

For such potential xml-model processing instructions, xml-model processors shall report to the application the *parsing result* of invoking the *rules for parsing pseudo-attributes from a string*, using the processing instruction information item's [content] property as the string.

A potential xml-model processing instruction is said to be an *xml-model processing instruction* if the *parsing result* is not an error when invoking the *rules for parsing pseudo-attributes from a string*, using the processing instruction information item's [content] property as the string.

Documents shall not contain processing instruction information items with the [target] property xml-model that are not xml-model processing instructions.

An xml-model processor shall process all xml-model processing instructions properly and shall pass on to the application the full *parsing result* for each xml-model processing instruction.

An xml-model processing instruction shall match the following production:

```
XmlModelPI ::= "<?xml-model" (S PseudoAtts)? - (Char* "?>" Char*) "?>"
```

Documents may specify the following *pseudo-attributes* on xml-model processing instructions, unless otherwise stated:

href
Specifies the location of the referenced schema. Documents shall specify this *pseudo-attribute*. Documents shall set the *value* to a string that matches the grammar for <IRI-reference> given in RFC 3987 *[RFC3987]*.

type
Specifies the content type of the referenced schema. If unspecified, the xml-model processor should return a *parsing result* that would be identical to that when the *value* is given as application/xml. The *value* of this *pseudo-attribute* is advisory in that it is intended to be used by an application only when no other source of media type information becomes available during retrieval of the schema itself.

schematypens
Specifies the namespace name of the schema language in which the referenced schema is written. The application can use this value when determining whether it can make use of the referenced schema.

charset
Specifies the character encoding for the referenced schema. If specified, documents shall set the *value* to a valid character encoding name, which shall be the name or alias labeled as "preferred MIME name" in the IANA Character Sets registry, if there is one, or the encoding's name, if none of the aliases are so labeled *[IANACHARSET]*.

title
Gives the title (or other human readable description) of the referenced schema. If specified, documents may use any string as the *value*.

group
If, for any xml-model processing instruction, its group *pseudo-attribute* has a non-empty *value*, special rules for associating schemas apply as follows:

1. By default only schemas which do not have a group *pseudo-attribute* specified or schemas which have an empty value in the group *pseudo-attribute* on the corresponding xml-model processing instruction are treated as being associated with XML document.

2. An application may provide an interface for specifying a group name. If the group name is specified, only schemas which have the same value specified in the group *pseudo-attribute* on the corresponding xml-model processing instruction are considered to be associated with the XML document.

phase
Gives the phase name of the validation function for use with a Schematron schema. If specified, documents may use any string as the *value*. If specified, the xml-model processor should include this information in the *parsing result* (regardless of the language of the associated schema).

If the associated schema is a Schematron schema, and the *parsing result* includes the phase *pseudo-attribute*, then the application is expected to use the *value* of this *pseudo-attribute* as the phase name of the validation function (see ISO/IEC 19757-3:2006, 6.1 *[ISO/IEC 19757-3]*).

To allow for extensibility, documents may specify other *pseudo-attributes* on xml-model processing instructions.

This part of ISO/IEC 19757 provides a way to associate multiple schemas with a given XML document. Furthermore, there exist other ways certain schemas can be associated with a given XML document. Regardless of the association method, this part of ISO/IEC 19757 does not prescribe the processing order when multiple schemas are associated with a given XML document.

In particular, this part of ISO/IEC 19757 does not define the interaction of xml-model processing instructions with xsi:schemaLocation and xsi:noNamespaceSchemaLocation attributes which provide hints for locating schema in W3C XML Schema. Applications supporting both xml-model processing instructions and xsi:schemaLocation/xsi:noNamespaceSchemaLocation attributes may provide means for specifying which information takes precedence.