
**Information technology — Document
Schema Definition Languages
(DSDL) —**

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

iTeh STANDARD PREVIEW

Technologies de l'information — Langages de définition de schéma de documents (DSDL) —

Partie 3: Validation basée sur des règles à l'aide de Schematron

<https://standards.iteh.ai/catalog/standards/sist/2da3fd10-f8d1-4a6d-bc1a-a8222eacea71/iso-iec-19757-3-2020>



iTeh STANDARD PREVIEW
(standards.iteh.ai)

[ISO/IEC 19757-3:2020](https://standards.iteh.ai/catalog/standards/sist/2da3fd10-f8d1-4a6d-bc1a-a8222eacea71/iso-iec-19757-3-2020)

<https://standards.iteh.ai/catalog/standards/sist/2da3fd10-f8d1-4a6d-bc1a-a8222eacea71/iso-iec-19757-3-2020>



COPYRIGHT PROTECTED DOCUMENT

© ISO/IEC 2020

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
Fax: +41 22 749 09 47
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Contents

	Page
Foreword	v
Introduction	vi
1 Scope	1
2 Normative references	1
3 Terms and definitions	1
4 Notation	4
5 Syntax	4
5.1 Well-formedness	4
5.2 Namespace	4
5.3 Whitespace	4
5.4 Core elements	4
5.4.1 General	4
5.4.2 active element	4
5.4.3 assert element	4
5.4.4 extends element	5
5.4.5 include element	5
5.4.6 let element	5
5.4.7 name element	5
5.4.8 ns element	6
5.4.9 param element	6
5.4.10 pattern element	6
5.4.11 phase element	6
5.4.12 report element	6
5.4.13 rule element	7
5.4.14 schema element	7
5.4.15 value-of element	7
5.5 Ancillary elements and attributes	8
5.5.1 diagnostic element	8
5.5.2 diagnostics element	8
5.5.3 dir element	8
5.5.4 emph element	8
5.5.5 flag attribute	8
5.5.6 fpi attribute	8
5.5.7 icon attribute	8
5.5.8 p element	8
5.5.9 properties element	8
5.5.10 property element	8
5.5.11 role attribute	9
5.5.12 see attribute	9
5.5.13 span element	9
5.5.14 subject attribute	9
5.5.15 title element	9
6 Semantics	9
6.1 Validation function	9
6.2 Minimal syntax (informative)	10
6.3 Abstract pattern processing	11
6.4 Query language binding	12
6.5 Order and side-effects	13
7 Conformance	13
7.1 Simple conformance	13
7.2 Full conformance	14

Annex A (normative) RELAX NG schema for Schematron	15
Annex B (normative) Schematron schema for additional constraints	19
Annex C (normative) Default query language binding	20
Annex D (informative) Schematron Validation Report Language	21
Annex E (informative) Design requirements	26
Annex F (informative) Use of Schematron as a vocabulary	27
Annex G (informative) Use of Schematron for multi-lingual schemas	28
Annex H (normative) Query language binding for XSLT2	29
Annex I (normative) Query language binding for XPath2	31
Annex J (normative) Query language binding for XSLT3	32
Annex K (normative) Query language binding for XPath3	34
Annex L (informative) Query language binding for EXSLT	35
Annex M (informative) Query language binding for STX	36
Annex N (informative) Example usage of Schematron properties	37
Bibliography	39

iTeh STANDARD PREVIEW
(standards.iteh.ai)

[ISO/IEC 19757-3:2020](https://standards.iteh.ai/catalog/standards/sist/2da3fd10-f8d1-4a6d-bc1a-a8222ecee71/iso-iec-19757-3-2020)

<https://standards.iteh.ai/catalog/standards/sist/2da3fd10-f8d1-4a6d-bc1a-a8222ecee71/iso-iec-19757-3-2020>

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).

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 <http://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.

This document was prepared by Joint Technical Committee 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 19757-3:2016), which has been technically revised.

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

- query language bindings have been added for XSLT 3.0 ([Annex J](#)) and XPath 3.0 ([Annex K](#));
- annexes pertaining to XPath and XSLT query language bindings ([Annexes H to K](#)) are now all normative, while those for EXSLT ([Annex L](#)) and STX ([Annex M](#)) remain informative.

A list of all parts in the ISO/IEC 19757 series can be found on the ISO website.

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.

Introduction

ISO/IEC 19757 (all parts) defines a set of Document Schema Definition Languages (DSDL) that can be used to specify one or more validation processes performed against Extensible Markup Language (XML) or Standard Generalized Markup Language (SGML) documents. [XML is an application profile SGML (see ISO 8879).]

A document model is an expression of the constraints to be placed on the structure and content of documents to be validated with the model. A number of technologies have been developed through various formal and informal consortia since the development of Document Type Definitions (DTDs) as part of ISO 8879, notably by the World Wide Web Consortium (W3C) and the Organization for the Advancement of Structured Information Standards (OASIS). A number of validation technologies are standardized in DSDL to complement those already available as standards or from the industry.

Through the validation that a structured document conforms to specified constraints in structure and content, the potentially many applications acting on the document are relieved from duplicating the task of confirming that such requirements have been met. Historically, such tasks and expressions have been developed and utilized in isolation, without consideration of how the features and functionality available in other technologies can enhance validation objectives.

The main objective of ISO/IEC 19757 (all parts) is to bring together different validation-related tasks and expressions to form a single extensible framework that allows technologies to work in series or in parallel to produce a single or a set of validation results. The extensibility of DSDL accommodates validation technologies not yet designed or specified.

In the past, different design and use criteria have led users to choose different validation technologies for different portions of their information. Bringing together information within a single XML document sometimes prevents existing document models from being used to validate sections of data. By providing an integrated suite of constraint description languages that can be applied to different subsets of a single XML document, ISO/IEC 19757 (all parts) allows different validation technologies to be integrated under a well-defined validation policy.

The structure of this document is as follows. [Clause 5](#) describes the syntax of an ISO Schematron schema. [Clause 6](#) describes the semantics of a correct ISO Schematron schema; the semantics specify when a document is valid with respect to an ISO Schematron schema. [Clause 7](#) describes conformance requirements for implementations of ISO Schematron validators. Annex A provides the ISO/IEC 19757-2 (RELAX NG) schema for ISO Schematron. [Annex B](#) provides the ISO Schematron schema for constraints in ISO Schematron that cannot be expressed by the schema of [Annex A](#). [Annex C](#) provides the default query language binding to XSLT1. Annex D provides an ISO/IEC 19757-2 (RELAX NG compact syntax) schema and corresponding ISO Schematron schema for a simple XML language Schematron Validation Report Language. [Annex E](#) provides motivating design requirements for ISO Schematron. [Annex F](#) specifies certain Schematron elements to be used in external vocabularies. [Annex G](#) provides a simple example of a multi-lingual schema. [Annexes H to M](#) provide query language bindings. [Annex N](#) shows example usage of Schematron properties.

This edition is backwards compatible with ISO/IEC 19757-3:2016, supersedes it and provides extra query language bindings, in particular for XSLT3.

Considered as a document type, a Schematron schema contains natural-language assertions concerning a set of documents, marked up with various elements and attributes for testing these natural-language assertions and for simplifying and grouping assertions.

Considered theoretically, a Schematron schema reduces to a non-chaining rule system whose terms are Boolean functions invoking an external query language on the instance and other visible XML documents, with syntactic features to reduce specification size and to allow efficient implementation.

Considered analytically, Schematron has two characteristic high-level abstractions: the pattern and the phase. These allow the representation of non-regular, non-sequential constraints that ISO/IEC 19757-2 cannot specify and various dynamic or contingent constraints.

This document is based on the Schematron^[2] assertion language. The `let` element is based on XCSL^[4]. Other features arise from the half-dozen early open-source implementations of Schematron in diverse programming languages and from discussions in electronic forums by Schematron users and implementers.

iTeh STANDARD PREVIEW (standards.iteh.ai)

[ISO/IEC 19757-3:2020](https://standards.iteh.ai/catalog/standards/sist/2da3fd10-f8d1-4a6d-bc1a-a8222ecea71/iso-iec-19757-3-2020)

<https://standards.iteh.ai/catalog/standards/sist/2da3fd10-f8d1-4a6d-bc1a-a8222ecea71/iso-iec-19757-3-2020>

iTeh STANDARD PREVIEW
(standards.iteh.ai)

[ISO/IEC 19757-3:2020](https://standards.iteh.ai/catalog/standards/sist/2da3fd10-f8d1-4a6d-bc1a-a8222ecea71/iso-iec-19757-3-2020)

<https://standards.iteh.ai/catalog/standards/sist/2da3fd10-f8d1-4a6d-bc1a-a8222ecea71/iso-iec-19757-3-2020>

Information technology — Document Schema Definition Languages (DSDL) —

Part 3: Rule-based validation using Schematron

1 Scope

This document specifies Schematron, a schema language for XML. This document establishes requirements for Schematron schemas and specifies when an XML document matches the patterns specified by a Schematron schema. Schematron uses query languages such as XPath for writing assertions.

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.

XPath¹⁾, *XML Path Language (XPath) Version 1.0*, W3C Recommendation, 16 November 1999

XPath2²⁾, *XML Path Language (XPath) 2.0*, W3C Recommendation, 23 January 2007

XPath3³⁾, *XML Path Language (XPath) 3.0*, W3C Recommendation, 8 April 2014

XPath2 Functions⁴⁾, *XQuery 1.0 and XPath 2.0 Functions and Operators*, W3C Recommendation, 23 January 2007

XPath3 Functions⁵⁾, *XPath and XQuery Functions and Operators 3.0*, W3C Recommendation, 8 April 2014

XSLT1⁶⁾, *XSL Transformations (XSLT) Version 1.0*, W3C Recommendation, 16 November 1999

XSLT2⁷⁾, *XSL Transformations (XSLT) Version 2.0*, W3C Recommendation, 23 January 2007

XSLT3⁸⁾, *XSL Transformations (XSLT) Version 3.0*, W3C Recommendation, 8 June 2017

ISO/IEC 19757-2, *Information technology — Document Schema Definition Language (DSDL) — Part 2: Regular-grammar-based validation — RELAX NG*

3 Terms and definitions

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

- 1) Available at <http://www.w3.org/TR/xpath>.
- 2) Available at <http://www.w3.org/TR/xpath20/>.
- 3) Available at <https://www.w3.org/TR/xpath-30/>.
- 4) Available at <http://www.w3.org/TR/xpath-functions/>.
- 5) Available at <https://www.w3.org/TR/xpath-functions-30/>.
- 6) Available at <http://www.w3.org/TR/xslt>.
- 7) Available at <http://www.w3.org/TR/xslt20/>.
- 8) Available at <https://www.w3.org/TR/xslt-30/>.

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

- ISO Online browsing platform: available at <https://www.iso.org/obp>
- IEC Electropedia: available at <http://www.electropedia.org/>

3.1 abstract pattern

pattern (3.13) in a *rule* (3.18) that has been parameterized to enable reuse

3.2 abstract rule

collection of *assertions* (3.5) which can be included in other *rules* (3.18) but which does not fire itself

3.3 active pattern

pattern (3.13) belonging to the *active phase* (3.4)

3.4 active phase

phase (3.14) whose *patterns* (3.13) are used for validation

3.5 assertion

natural-language statement with associated *assertion test* (3.6) and ancillary attributes

3.6 assertion test

Boolean query

iTeh STANDARD PREVIEW
(standards.iteh.ai)

Note 1 to entry: An assertion test "succeeds" or "fails".

ISO/IEC 19757-3:2020

<https://standards.iteh.ai/catalog/standards/sist/2da3fd10-f8d1-4a6d-bc1a-a8222ecee71/iso-iec-19757-3-2020>

3.7 Schematron schema

document that satisfies all the requirements of this document

3.8 diagnostic

named natural language statements providing information to end-users of validators concerning the expected and actual values together with repair hints

3.9 elaborated rule-context expression

single *rule-context expression* (3.20) which explicitly disallows items selected by lexically previous *rule contexts* (3.19) in the same *pattern* (3.13)

3.10 implementation

implementation of a Schematron validator

3.11 name

mixture of name characters with a restricted set of initial characters

Note 1 to entry: See Production 5 of XML1.

3.12 natural-language assertion

natural-language statement expressing some part of a *pattern* (3.13)

Note 1 to entry: A natural-language assertion is "met" or "unmet".

3.13**pattern**

unordered collection of *rules* (3.18) with an optional identifier and ancillary information

3.14**phase**

named, unordered collection of *patterns* (3.13)

Note 1 to entry: Patterns may belong to more than one phase.

Note 2 to entry: Two strings, #ALL and #DEFAULT, are reserved with particular meanings.

3.15**progressive validation**

validation of constraints in stages determined or grouped to some extent by the schema author rather than, for example, entirely determined by the document order

3.16**property**

named data giving additional metadata on an *assertion* (3.5) or report

3.17**query language binding**

named set, specified in a document called a Query Language Binding, of the languages and conventions used for *assertion tests* (3.6), *rule-context expressions* (3.20) and so on, by a particular Schematron implementation (3.10)

iTeh STANDARD PREVIEW

3.18**rule**

unordered collection of *assertions* (3.5) with a *rule-context expression* (3.20) and ancillary attributes

ISO/IEC 19757-3:2020

3.19**rule context**

element or other information item used for *assertion tests* (3.6)

<https://standards.iteh.ai/catalog/standards/sist/2da3fd10-f8d1-4a6d-bc1a-a8222ecee71/iso-iec-19757-3-2020>

Note 1 to entry: A rule is said to fire when an information item matches the rule context.

3.20**rule-context expression**

query to specify *subjects* (3.21)

Note 1 to entry: A *rule context* (3.19) is said to match an information item when that information item has not been matched by any lexically-previous rule-context expressions in the same *pattern* (3.13) and the information item is one of the information items that the query would specify.

3.21**subject**

particular information item which corresponds to the object of interest of the *natural-language assertions* (3.12) and typically is matched by the context expression of a *rule* (3.18)

3.22**valid**

<schema> passing all *assertion tests* (3.6) in fired *rules* (3.18) of *active patterns* (3.3)

3.23**variable**

constant value, represented by a *name* (3.11), evaluated within the parent schema, *phase* (3.14), *pattern* (3.13) or *rule* (3.18) and scoped within the parent schema, phase, pattern or rule

4 Notation

This document uses XPath to identify information items, to the extent that items in the XPath data model can be derived from those defined by XML Infoset, in Schematron schemas.

5 Syntax

5.1 Well-formedness

A Schematron schema shall be a well-formed XML document, according to the version of XML used.

5.2 Namespace

All elements shown in the grammar for Schematron are qualified with the namespace URI [IRI]:

<http://purl.oclc.org/dsdl/schematron>

In subsequent clauses, the prefix `sch` is taken as bound to the Schematron namespace URI for exposition purposes. The prefix `sch` is not reserved or required by this document. Any element can also have foreign attributes in addition to the attributes shown in the grammar. A foreign attribute is an attribute with a name whose namespace URI is neither the empty string nor the Schematron namespace URI. Any non-empty element may have foreign child elements in addition to the child elements shown in the grammar. A foreign element is an element with a name whose namespace URI is not the Schematron namespace URI. There are no constraints on the relative position of foreign child elements with respect to other child elements.

5.3 Whitespace

Any element can also have children strings that consist entirely of whitespace characters, where a whitespace character is one of U+0020, U+009, U+00D or U+00A. There are no constraints on the relative position of whitespace string children with respect to child elements.

Leading and trailing whitespace should be stripped from attributes defined by this document. Whitespace should be collapsed in elements defined by this document that allow text. Whitespace may be stripped from elements defined by this document that do not allow text.

5.4 Core elements

5.4.1 General

A Schematron schema shall follow the grammar given in [Annex A](#).

For information on the use of Schematron elements in external vocabularies, see [Annex F](#).

5.4.2 active element

The required `pattern` attribute is a reference to a pattern that is active in the current phase.

5.4.3 assert element

This element indicates an assertion made about the context nodes. The data content is a natural-language assertion. The required `test` attribute is an assertion test evaluated in the current context. If the test evaluates positive, the assertion succeeds. The optional `diagnostics` attribute is a reference to further diagnostic information.

The natural-language assertion shall be a positive statement of a constraint.

The `icon`, `see` and `fpi` attributes ([5.5.7](#), [5.5.12](#) and [5.5.6](#)) allow rich interfaces and documentation.

The `flag` attribute (5.5.5) allows more detailed outcomes.

The `role` and `subject` attributes (5.5.11 and 5.5.14) allow explicit identification of some part of a pattern.

The natural-language assertion may contain information about actual values in addition to expected values and may contain diagnostic information. Users should note, however, that the `diagnostic` element is provided for such information to encourage clear statement of the natural-language assertion.

5.4.4 `extends` element

The `extends` element allows reference to the contents of other declarations. The `extends` element shall either have an `href` attribute or a `rule` attribute but not both.

Abstract rules are named lists of assertions without a context expression. An `extends` element with a `rule` attribute shall reference an abstract rule. The current rule uses all the assertions from the abstract rule it extends.

An `extends` element with an `href` attribute shall reference external declarations. The `href` attribute is an IRI reference to an external well-formed XML document or to an element in an external well-formed XML document that is a Schematron element of the same type as the parent element of the `extends` element. The contents of that referenced element shall be inserted in place of the `extends` element.

In such a case, the relative position of elements in the post-inclusion document may be to that extent invalid against the schema for Schematron in Annex A; however, other schema constraints such as containment shall still apply.

5.4.5 `include` element

The required `href` attribute shall be an IRI reference to a well-formed XML document or to an element in a well-formed XML document.

The referenced element shall be inserted in place of the `include` element. The referenced element shall be a type which is allowed by the grammar for Schematron at the location of the `include` element.

5.4.6 `let` element

The `let` element declares a named variable. If the `let` element is the child of a `rule` element, the variable is calculated and scoped to the current rule and context. Otherwise, the variable is calculated with the context of the instance document root.

The required `name` attribute is the name of the variable. The `value` attribute is an expression evaluated in the current context. If no `value` attribute is specified, the value of the attribute is the element content of the `let` element.

It is an error to reference a variable that has not been defined in the current schema, phase, pattern or rule, if the query language binding allows this to be determined reliably. It is an error for a variable to be multiply defined in the current schema, phase, pattern and rule.

The variable is substituted into assertion tests and other expressions in the same rule before the test or expression is evaluated. The query language binding specifies which lexical conventions are used to detect references to variables.

5.4.7 `name` element

The `name` element provides the names of nodes from the instance document to allow clearer assertions and diagnostics. The optional `path` attribute is an expression evaluated in the current context that returns a string that is the name of a node. If the `path` attribute is omitted, the name of the context node is used.