BS ISO/IEC 19757-3:2016

BSI Standards Publication

Information technology — Document Schema Definition Languages (DSDL)

Part 3: Rule-based validation — Schematron

bsi.

...making excellence a habit.™

Document Sharing Hub

Full document details: BS ISO/IEC 19757-3:2016

National foreword

This British Standard is the UK implementation of ISO/IEC 19757-3:2016. It supersedes BS ISO/IEC 19757-3:2006 which is withdrawn.

The UK participation in its preparation was entrusted to Technical Committee IST/41, Document description and processing language.

A list of organizations represented on this committee can be obtained on request to its secretary.

This publication does not purport to include all the necessary provisions of a contract. Users are responsible for its correct application.

© The British Standards Institution 2016.
Published by BSI Standards Limited 2016

ISBN 978 0 580 71314 9

ICS 35.240.30

Compliance with a British Standard cannot confer immunity from legal obligations.

This British Standard was published under the authority of the Standards Policy and Strategy Committee on 31 January 2016.

Amendments/corrigenda issued since publication

Date	Text affected

Documents
Document Sharing Hub

Full document details: BS ISO/IEC 19757-3:2016

INTERNATIONAL STANDARD

ISO/IEC 19757-3

Second edition
2016-01-15

Information technology — Document Schema Definition Languages (DSDL) —

Part 3: Rule-based validation — Schematron

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

Partie 3: Validation de règles orientées — Schematron

Reference number
ISO/IEC 19757-3:2016(E)

Documents
Document Sharing Hub
© ISO/IEC 2016

Full document details: BS ISO/IEC 19757-3:2016

COPYRIGHT PROTECTED DOCUMENT

© ISO/IEC 2016, Published in Switzerland

All rights reserved. Unless otherwise specified, 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
Ch. de Blandonnet 8 • CP 401
CH-1214 Vernier, Geneva, Switzerland
Tel. +41 22 749 01 11
Fax +41 22 749 09 47
copyright@iso.org
www.iso.org

© ISO/IEC 2016 – All rights reserved

Documents
Document Sharing Hub

Full document details: BS ISO/IEC 19757-3:2016

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 `active` element ... 4
5.4.2 `assert` element ... 4
5.4.3 `extends` element ... 5
5.4.4 `include` element ... 5
5.4.5 `let` element ... 5
5.4.6 `name` element ... 5
5.4.7 `ns` element ... 6
5.4.8 `param` element ... 6
5.4.9 `pattern` element ... 6
5.4.10 `phase` element ... 6
5.4.11 `report` element ... 6
5.4.12 `rule` element ... 7
5.4.13 `schema` element ... 7
5.4.14 `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 ... 9
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 ... 10
6.3 Abstract pattern processing ... 10
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 ... 13
Annex A (normative) **RELAX NG schema for Schematron** ... **15**

AVA Documents
Document Sharing Hub

© ISO/IEC 2016 – All rights reserved

Full document details: BS ISO/IEC 19757-3:2016

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 XSLT 2** **29**
Annex I (informative) **Query language binding for XPath 2** **31**
Annex J (informative) **Query language binding for EXSLT** **32**
Annex K (informative) **Query language binding for STX** **33**
Annex L (informative) **Example usage of Schematron properties** **34**
Bibliography **36**

© ISO/IEC 2016 – All rights reserved
Documents
Document Sharing Hub
Full document details: BS ISO/IEC 19757-3:2016

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.

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).

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

For an explanation on the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the WTO principles in the Technical Barriers to Trade (TBT) see the following URL: Foreword - Supplementary information

The committee responsible for this document is ISO/IEC JTC 1, *Information technology*, Subcommittee SC 34, *Document description and processing languages.*

This second edition cancels and replaces the first edition (ISO/IEC 19757-3:2006), which has been technically revised.

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

— *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*

© ISO/IEC 2016 – All rights reserved

Documents
Document Sharing Hub
Full document details: BS ISO/IEC 19757-3:2016

Introduction

ISO/IEC 19757 defines a set of Document Schema Definition Languages (DSDL) that can be used to specify one or more validation processes performed against Extensible Stylesheet 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 industry.

To validate that a structured document conforms to specified constraints in structure and content relieves the potentially many applications acting on the document from having to duplicate 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 might enhance validation objectives.

The main objective of ISO/IEC 19757 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 allows different validation technologies to be integrated under a well-defined validation policy.

The structure of this part of ISO/IEC 19757 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 is a normative annex providing the ISO/IEC 19757-2 (RELAX NG) schema for ISO Schematron. Annex B is a normative annex providing the ISO Schematron schema for constraints in ISO Schematron that cannot be expressed by the schema of Annex A. Annex C is a normative annex providing the default query language binding to XSLT1. Annex D is a non-normative annex providing 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 is a non-normative annex providing motivating design requirements for ISO Schematron. Annex F is a normative annex allowing certain Schematron elements to be used in external vocabularies. Annex G is a non-normative annex with a simple example of a multi-lingual schema.

This edition is backwards compatible with ISO/IEC 19757-3:2006, supercedes it, and augments it with the following capabilities: patterns may validate different documents, the inclusion mechanism has been supplemented by an enhanced extension mechanism, assertions may have linked properties, and SVRL may take richer text. As well, this edition provides extra query language bindings, in particular for XSLT2.

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.

© ISO/IEC 2016 – All rights reserved
Documents
Document Sharing Hub
Full document details: BS ISO/IEC 19757-3:2016

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 part of ISO/IEC 19757 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.

© ISO/IEC 2016 – All rights reserved

Documents
Document Sharing Hub
Full document details: BS ISO/IEC 19757-3:2016

Documents
Document Sharing Hub

Full document details: BS ISO/IEC 19757-3:2016

INTERNATIONAL STANDARD ISO/IEC 19757-3:2016(E)

Information technology — Document Schema Definition Languages (DSDL) —

Part 3: Rule-based validation — Schematron

1 Scope

This part of ISO/IEC 19757 specifies Schematron, a schema language for XML. This part of ISO/IEC 19757 establishes requirements for Schematron schemas and specifies when an XML document matches the patterns specified by a Schematron schema.

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.

NOTE Each of the following documents has a unique identifier that is used to cite the document in the text. The unique identifier consists of the part of the reference up to the first comma.

XPath, *XML Path Language (XPath) Version 1.0*, W3C Recommendation, 16 November 1999,[1)]

XPath2, *XML Path Language (XPath) 2.0*, W3C Recommendation, 23 January 2007,[2)]

XPath2 Functions, *XQuery 1.0 and XPath 2.0 Functions and Operators*, W3C Recommendation, 23 January 2007,[3)]

XSLT1, *XSL Transformations (XSLT) Version 1.0*, W3C Recommendation, 16 November 1999,[4)]

XSLT2, *XSL Transformations (XSLT) Version 2.0*, W3C Recommendation, 23 January 2007,[5)]

3 Terms and definitions

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

3.1
abstract pattern
pattern in a rule that has been parameterized to enable reuse

3.2
abstract rule
collection of assertions which can be included in other rules but which does not fire itself

1) http://www.w3.org/TR/xpath

2) http://www.w3.org/TR/xpath20/

3) http://www.w3.org/TR/xpath-functions/

4) http://www.w3.org/TR/xslt

5) http://www.w3.org/TR/xslt20/

© ISO/IEC 2016 – All rights reserved

Documents
Document Sharing Hub

Full document details: BS ISO/IEC 19757-3:2016

3.3
active pattern
pattern belonging to the active phase

3.4
active phase
one particular phase, whose patterns are used for validation

3.5
assertion
natural-language assertion with corresponding assertion test and ancillary attributes: assertions are marked up with assert and report elements

3.6
assertion test
assertion modelled or implemented by a Boolean query

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

3.7
compound document
notional instance document which has been divided into the original instance and a number of subordinate well-formed XML documents

Note 1 to entry: These subordinate documents may also contain schema-like information.

3.8
correct schema
schema that satisfies all the requirements of this part of ISO/IEC 19757

3.9
diagnostic
named natural language statements providing information to end-users of validators concerning the expected and actual values together with repair hints

3.10
elaborated rule context expression
single rule context expression which explicitly disallows items selected by lexically previous rule contexts in the same pattern

3.11
good schema
correct schema with queries which terminate and do not add constraints to those of the natural-language assertions

Note 1 to entry: It may not be possible to compute that a schema is good.

3.12
implementation
implementation of a Schematron validator

3.13
name
token with no whitespace characters

3.14
natural-language assertion
natural-language statement expressing some part of a pattern

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

© ISO/IEC 2016 – All rights reserved
Documents
Document Sharing Hub
Full document details: BS ISO/IEC 19757-3:2016

3.15
pattern
named structure in instances specified in a schema by a lexically-ordered collection of rules

3.16
phase
named, unordered collection of patterns; patterns may belong to more than one phase; two names, #ALL and #DEFAULT, are reserved with particular meanings

3.17
progressive validation
validation of constraints in stages determined or grouped to some extent by the schema author rather than, for example, entirely determined by document order

3.18
property
named data giving additional metadata on an assertion or report

3.19
query language binding
named set, specified in a document called a Query Language Binding, of the languages and conventions used for assertion tests, rule-context expressions and so on, by a particular Schematron implementation

3.20
rule
unordered collection of assertions with a rule-context expression and ancillary attributes

3.21
rule context
element or other information item used for assertion tests

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

3.22
rule-context expression
query to specify subjects

Note 1 to entry: A rule-context 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 and the information item is one of the information items that the query would specify.

3.23
schema
specification of a set of XML documents

3.24
subject
particular information item which corresponds to the object of interest of the natural-language assertions and typically is matched by the context expression of a rule

3.25
valid with respect to a schema
member of the set of XML documents described by the schema: an instance document is valid if no assertion tests in fired rules of active patterns fail

3.26
variable
constant value, evaluated within the parent schema, phase, pattern or rule and scoped within the parent schema, phase, pattern or rule

© ISO/IEC 2016 – All rights reserved

Documents
Document Sharing Hub
Full document details: BS ISO/IEC 19757-3:2016

4 Notation

This part of ISO/IEC 19757 uses XPath to identify information items 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 part of ISO/IEC 19757. 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 as 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.

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

5.4 Core elements

The grammar for Schematron elements is given in Annex A.

5.4.1 `active` element

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

5.4.2 `assert` element

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 report 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 allow rich interfaces and documentation. They are defined below.

The `flag` attribute allows more detailed outcomes. It is defined below.

© ISO/IEC 2016 – All rights reserved
Documents
Document Sharing Hub
Full document details: BS ISO/IEC 19757-3:2016

The `role` and `subject` attributes allow explicit identification of some part of a pattern. They are defined below.

NOTE 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.

5.4.3 `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 attribute 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 external well-formed XML document that is 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 Annex A; however, other schema constraints such as containment shall still apply.

NOTE The capability to extend with multiple elements was not part of ISO/IEC 19757-3:2006.

5.4.4 `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.

5.4.5 `let` element

A declaration of 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.

5.4.6 `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. In the latter case, the name of the node is used.

© ISO/IEC 2016 – All rights reserved

Documents
Document Sharing Hub
Full document details: BS ISO/IEC 19757-3:2016