INTERNATIONAL STANDARD

ISO/IEC 19757-9

First edition
2008-11-15

Information technology — Document Schema Definition Languages (DSDL) —

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

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

Partie 9: Déclaration d'espace de nommage et de type de données dans les définitions de type de document (DTD)

Reference number
ISO/IEC 19757-9:2008(E)

PDF disclaimer

This PDF file may contain embedded typefaces. In accordance with Adobe's licensing policy, this file may be printed or viewed but shall not be edited unless the typefaces which are embedded are licensed to and installed on the computer performing the editing. In downloading this file, parties accept therein the responsibility of not infringing Adobe's licensing policy. The ISO Central Secretariat accepts no liability in this area.

Adobe is a trademark of Adobe Systems Incorporated.

Details of the software products used to create this PDF file can be found in the General Info relative to the file; the PDF-creation parameters were optimized for printing. Every care has been taken to ensure that the file is suitable for use by ISO member bodies. In the unlikely event that a problem relating to it is found, please inform the Central Secretariat at the address given below.

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

Page

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 19757-9 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*

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

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

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

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

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

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

Introduction

The language of Document Type Definitions (DTDs) was the original schema language defined by W3C XML and was closely based upon the DTD language defined by ISO 8879:1986. For a variety of reasons, both technical and economic, many users of XML for document-centric applications, especially among those who were previously (and in some cases continue to be) users of SGML, still favour the use of DTD language for grammar-based schema definition in such applications.

It is important to provide users that have made a significant investment in DTDs with a migration path that will enable them to adopt ISO/IEC 19757 without having to translate all their existing DTDs to a different schema language, especially as this would oblige them to replace all systems that only work with DTDs, with all the expense and organizational upheavals thereby entailed. A sensible migration path should enable such users to continue to use DTDs for as much of the document validation process as can reasonably be managed, but also enable them to reap the benefits of using those parts of ISO/IEC 19757 that most obviously complement and extend the use of DTDs for validation purposes.

It is equally important that the migration path should enable users to continue to use legacy systems that are incapable of using any kind of extension to the DTD language, while at the same time introducing new systems that are equipped to use such extensions. The method of extension will therefore be such that DTDs with extended functionality are valid XML DTDs in accordance with W3C XML. It should remain possible to validate an instance *that does not make any use of the extended functionality* against a DTD that contains the extended functionality, using legacy system tools, and achieve the same result as would be achieved if the DTD did not contain the extended functionality.

The most significant validation tasks that cannot be performed using a DTD alone are

— validation of names with respect to namespaces,
— validation of data content and attribute values with respect to datatypes,
— rules-based validation.

Of these three, the last is made possible by implementation of ISO/IEC 19757-3. The first two tasks are currently only possible if ISO/IEC 19757-2 is implemented, but existing DTD users are unlikely to wish to maintain parallel RELAX NG and DTD schemas for each application simply as a means of supporting the use of namespaces and datatypes.

For these reasons this part of ISO/IEC 19757 addresses the extension of DTDs to support the declaration of namespaces and datatypes.

Information technology — Document Schema Definition Languages (DSDL) —

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

1 Scope

This part of ISO/IEC 19757 defines a language that is designed to extend the declarative functionality of an XML DTD to include

— declaring one or more namespaces to which some or all of the element and attribute names in a DTD belong,
— declaring constraints on the content of elements with content model ANY to contain elements whose names belong to one or more specified namespaces,
— declaring datatypes for elements that contain data content only and for attribute values.

Two alternative syntax bindings for this language are defined. The first syntax binding uses XML processing instructions and is designed to enable declarations in this language to be embedded within an XML DTD without invalidating the DTD or altering its interpretation so far as legacy DTD parsers are concerned. This first syntax also provides a means of associating a DTD with an external declarations subset containing declarations in either syntax. This syntax is defined in Clause 4 using the modified BNF syntax notation used in W3C XML.

The second syntax binding uses an XML document syntax and is defined in Clause 5. The syntax rules are defined by a schema that conforms to the RELAX NG Compact Syntax defined in ISO/IEC 19757-2. This syntax is designed to enable declarations in this language to be expressed almost entirely in XML (in this case one XML processing instruction needs to be inserted in the DTD), to facilitate implementation using existing XML tools, either as a namespace-qualified fragment embedded within an XML instance or as a separate XML document.

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.

ISO 8879:1986, *Information processing — Text and office systems — Standard Generalized Markup Language (SGML)*

W3C XML, *Extensible Markup Language (XML) 1.0 (Fourth Edition)*, W3C Recommendation, 16 August 2006, http://www.w3.org/TR/2006/REC-xml-20060816/

W3C XML-Names, *Namespaces in XML 1.0 (Second Edition)*, W3C Recommendation, 16 August 2006, http://www.w3.org/TR/2006/REC-xml-names-20060816/

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

IETF RFC 3987, *Internationalized Resource Identifiers (IRIs)*, January 2005, http://www.ietf.org/rfc/rfc3987