



**International
Standard**

ISO 21720

**XLIFF (XML Localization
Interchange File Format)**

*XLIFF (Format de fichier XML pour l'échange de données de
localisation)*

**Second edition
2024-07**

iTeh Standards
(<https://standards.iteh.ai>)
Document Preview

[ISO 21720:2024](https://standards.iteh.ai/catalog/standards/iso/51b1b925-05b2-4822-9658-6246a29c13dd/iso-21720-2024)

<https://standards.iteh.ai/catalog/standards/iso/51b1b925-05b2-4822-9658-6246a29c13dd/iso-21720-2024>

iTeh Standards
(<https://standards.iteh.ai>)
Document Preview

[ISO 21720:2024](https://standards.iteh.ai/catalog/standards/iso/51b1b925-05b2-4822-9658-6246a29c13dd/iso-21720-2024)

<https://standards.iteh.ai/catalog/standards/iso/51b1b925-05b2-4822-9658-6246a29c13dd/iso-21720-2024>



COPYRIGHT PROTECTED DOCUMENT

© ISO 2024

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

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

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 ISO document should be noted (see www.iso.org/directives).

ISO draws attention to the possibility that the implementation of this document may involve the use of (a) patent(s). ISO takes no position concerning the evidence, validity or applicability of any claimed patent rights in respect thereof. As of the date of publication of this document, ISO had not received notice of (a) patent(s) which may be required to implement this document. However, implementers are cautioned that this may not represent the latest information, which may be obtained from the patent database available at www.iso.org/patents. ISO shall not be held responsible for identifying any or all such patent rights.

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 OASIS (as XLIFF Version 2.1, February 2018) and drafted in accordance with its editorial rules. It was assigned to Technical Committee ISO/TC 37, *Language and terminology*, Subcommittee SC 5, *Translation, interpreting and related technology*, and adopted under the “fast-track procedure”.

This second edition cancels and replaces the first edition (ISO 21720:2017), which has been technically revised.

The main changes are as follows (see also Appendix C):

- Two major features are being added in XLIFF Version 2.1:
 - Advanced Validation methods;
 - Native Support for ITS 2.0.
- The Change Tracking Module was demoted to an extension to free hands of the TC and other implementers while working on a new version of the Change Tracking Module for XLIFF 2.2.

ISO 21720:2024(en)

- A major bug fix was performed on the core xsd. The core xsd now enforces the xs:language data type on the srcLang and trgLang attributes. It was critical to make this fix, because -- as per OASIS policy -- validation artifacts would prevail over the prose provisions that are correct in both XLIFF 2.1 and XLIFF 2.0.
- Also an erroneously omitted Constraint of the xml:lang attribute on the <source> element has been added/restored in the normative text.
- Apart from the five (5) major changes mentioned above, numerous editorial bugfixes were made to secure greater clarity, either by fixing example errors or omissions, or by reorganizing normative content, so that the intent becomes clear and unequivocal at some troublesome places highlighted by XLIFF 2.0 implementers.
- Importantly, the TC decided to drop informative listings of the validation artifacts that had bloated the spec extent unnecessarily, were hard to keep in sync with the actual normative artifacts, while their actual usability proved rather limited -- readers who were able to read schema languages would not actually read them as printed listings and would anyways refer to the actual validation artifacts that are now referenced more prominently.

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.

iTeh Standards (<https://standards.iteh.ai>) Document Preview

[ISO 21720:2024](#)

<https://standards.iteh.ai/catalog/standards/iso/51b1b925-05b2-4822-9658-6246a29c13dd/iso-21720-2024>

Table of Contents

1 Introduction

- 1.0 IPR Policy
- 1.1 Terminology
 - 1.1.1 Key words
 - 1.1.2 Definitions
 - 1.1.3 Key concepts
- 1.2 Normative References
- 1.3 Non-Normative References

2 Conformance

3 Fragment Identification

- 3.1 Selectors for Core Elements
- 3.2 Selectors for Modules and Extensions
- 3.3 Relative References
- 3.4 Examples

4 The Core Specification

- 4.1 General Processing Requirements
- 4.2 Elements
 - 4.2.1 Tree Structure
 - 4.2.2 Structural Elements
 - 4.2.3 Inline Elements
- 4.3 Attributes
 - 4.3.1 XLIFF Attributes
 - 4.3.2 XML namespace
- 4.4 CDATA sections
- 4.5 XML Comments
- 4.6 XML Processing Instructions
- 4.7 Inline Content
 - 4.7.1 Text
 - 4.7.2 Inline Codes
 - 4.7.3 Annotations
 - 4.7.4 Sub-Flows
 - 4.7.5 White Spaces
 - 4.7.6 Bidirectional Text
 - 4.7.7 Target Content Modification
 - 4.7.8 Content Comparison
- 4.8 Segmentation
 - 4.8.1 Segments Representation
 - 4.8.2 Segments Order
 - 4.8.3 Segmentation Modification
 - 4.8.4 Best Practice for *Mergers* (Informative)
- 4.9 Extension Mechanisms
 - 4.9.1 Extension Points
 - 4.9.2 Constraints
 - 4.9.3 Processing Requirements

5 The Modules Specifications

- 5.1 Translation Candidates Module
 - 5.1.1 Introduction
 - 5.1.2 Module Namespace and Validation Artifacts
 - 5.1.3 Module Fragment Identification Prefix
 - 5.1.4 XLIFF Core Reuse
 - 5.1.5 Translation Candidate Annotation
 - 5.1.6 Module Elements
 - 5.1.7 Module Attributes

- [5.1.8 Example](#)
- [5.2 Glossary Module](#)
 - [5.2.1 Introduction](#)
 - [5.2.2 Module Namespace and Validation Artifacts](#)
 - [5.2.3 Module Fragment Identification Prefix](#)
 - [5.2.4 Module Elements](#)
 - [5.2.5 Module Attributes](#)
 - [5.2.6 Example](#)
- [5.3 Format Style Module](#)
 - [5.3.1 Introduction](#)
 - [5.3.2 Module Namespace and Validation Artifacts](#)
 - [5.3.3 Module Fragment Identification Prefix](#)
 - [5.3.4 Module Specification](#)
 - [5.3.5 Module Attributes](#)
- [5.4 Metadata Module](#)
 - [5.4.1 Introduction](#)
 - [5.4.2 Module Namespace and Validation Artifacts](#)
 - [5.4.3 Module Fragment Identification Prefix](#)
 - [5.4.4 Module Elements](#)
 - [5.4.5 Module Attributes](#)
- [5.5 Resource Data Module](#)
 - [5.5.1 Introduction](#)
 - [5.5.2 Module Namespace and Validation Artifacts](#)
 - [5.5.3 Module Fragment Identification Prefix](#)
 - [5.5.4 Module Elements](#)
 - [5.5.5 Module Attributes](#)
 - [5.5.6 Examples](#)
- [5.6 Change Tracking Extension \(Informative\)](#)
 - [5.6.1 Introduction](#)
 - [5.6.2 Module Namespace and Validation Artifacts](#)
 - [5.6.3 Module Fragment Identification Prefix](#)
 - [5.6.4 Module Elements](#)
 - [5.6.5 Module Attributes](#)
 - [5.6.6 Example](#)
- [5.7 Size and Length Restriction Module](#)
 - [5.7.1 Introduction](#)
 - [5.7.2 Module Namespace and Validation Artifacts](#)
 - [5.7.3 Module Fragment Identification Prefix](#)
 - [5.7.4 Module Elements](#)
 - [5.7.5 Module Attributes](#)
 - [5.7.6 Standard profiles](#)
 - [5.7.7 Third party profiles](#)
 - [5.7.8 Conformance](#)
 - [5.7.9 Example](#)
- [5.8 Validation Module](#)
 - [5.8.1 Introduction](#)
 - [5.8.2 Module Namespace and Validation Artifacts](#)
 - [5.8.3 Module Fragment Identification Prefix](#)
 - [5.8.4 Module Elements](#)
 - [5.8.5 Module Attributes](#)
- [5.9 ITS Module](#)
 - [5.9.1 Introduction](#)
 - [5.9.2 Module Namespace and Validation Artifacts](#)
 - [5.9.3 Module Fragment Identification Prefix](#)
 - [5.9.4 Conformance to the ITS Module Specification](#)

- [5.9.5 ITS Tools Referencing](#)
- [5.9.6 ITS data categories defined in the ITS Module](#)
- [5.9.7 ITS data categories that have a partial overlap with XLIFF features](#)
- [5.9.8 ITS data categories available through XLIFF Core and other Modules](#)
- [5.9.9 ITS data categories not represented in XLIFF](#)
- [5.9.10 ITS Mapping Annotations](#)
- [5.9.11 Module Elements](#)
- [5.9.12 Module Attributes](#)
- [5.9.13 Example file](#)

Appendixes

[A Media Type Registration Template for XLIFF Version 2.0 and higher Versions](#)

- [A.1 Registration Template](#)
- [A.1.1 Detailed Security Considerations](#)
- [A.1.1.1 Privacy, trust and integrity](#)
- [A.1.1.2 Core](#)
- [A.1.1.3 Resource Data Module](#)
- [A.1.1.4 ITS Module](#)
- [A.1.1.5 Other potentially security sensitive constructs](#)

[B Machine Readable Validation Artifacts \(Informative\)](#)

- [B.1 XML Schemas Tree](#)
- [B.2 Support Schemas](#)

[C Specification Change Tracking \(Informative\)](#)

- [C.1 High Level Summary of Changes made in Comparison to XLIFF Version 2.0](#)
- [C.2 Tracking of changes made in response to Public Reviews](#)
- [C.2.1 Tracking of changes in response to the Public Review of the Candidate OASIS Standard 01](#)
- [C.2.2 Tracking of changes in response to the 4th Public Review](#)
- [C.2.3 Tracking of changes in response to the 3rd Public Review](#)
- [C.2.4 Tracking of changes in response to the 2nd Public Review](#)
- [C.2.5 Tracking of changes in response to the 1st Public Review](#)

[D Acknowledgements \(Informative\)](#)

1 Introduction

XLIFF is the *XML Localization Interchange File Format* designed by a group of multilingual content publishers, software providers, localization service providers, localization tools providers, and researchers. It is intended to give any multilingual content owner a single interchange file format that can be understood by any localization provider, using any conformant localization tool. While the primary focus is on being a lossless interchange format, usage of XLIFF as a processing format is neither encouraged nor discouraged or prohibited.

All text is normative unless otherwise labeled. The following common methods are used for labeling portions of this specification as informative and hence non-normative:

Appendices and sections marked as "(Informative)" or "Non-Normative" in title,
Notes (sections with the "Note" title),
Warnings (sections with the "Warning" title),
Examples (mainly example code listings, tree diagrams, but also any inline examples or illustrative exemplary lists in otherwise normative text),
Schema and other validation artifacts listings (the corresponding artifacts are normative, not their listings).

1.0 IPR Policy

This specification is provided under the [RF on RAND Terms](#) Mode of the [OASIS IPR Policy](#), the mode chosen when the Technical Committee was established. For information on whether any patents have been disclosed that may be essential to implementing this specification, and any offers of patent licensing terms, please refer to the Intellectual Property Rights section of the Technical Committee web page (<https://www.oasis-open.org/committees/xliff/ipr.php>).

1.1 Terminology

1.1.1 Key words

The key words MUST, MUST NOT, REQUIRED, SHALL, SHALL NOT, SHOULD, SHOULD NOT, RECOMMENDED, MAY, and OPTIONAL are to be interpreted as described in [RFC 2119](#).

1.1.2 Definitions

Agent

any application or tool that generates (creates), reads, edits, writes, processes, stores, renders or otherwise handles *XLIFF Documents*.

Agent is the most general application conformance target that subsumes all other specialized user agents disregarding whether they are defined in this specification or not.

Enrich, Enriching

the process of associating module and extension based metadata and resources with the *Extracted* XLIFF payload

Processing Requirements

- *Enriching* MAY happen at the time of *Extraction*.

Note

Extractor knowledge of the native format is not assumed while *Enriching*.

Enricher, Enricher Agent

any *Agent* that performs the *Enriching* process

Extract, Extraction

the process of encoding localizable content from a native content or User Interface format as XLIFF payload, so that localizable parts of the content in the source language are available for *Translation* into the target language along with the necessary context information

Extractor, Extractor Agent

any *Agent* that performs the *Extraction* process

Merge, Merging

the process of importing XLIFF payload back to the originating native format, based on the *full knowledge* of the *Extraction* mechanism, so that the localized content or User Interface strings replace the source language in the native format

Merger, Merger Agent

an *Agent* that performs the *Merge* process

Warning

Unless specified otherwise, any *Merger* is deemed to have the same knowledge of the native format as the *Extractor* throughout the specification.

Mergers independent of *Extractors* can succeed, but it is out of scope of this specification to specify interoperability for *Merging* back without the full *Extractor* knowledge of the native format.

Modify, Modification

the process of changing core and module XLIFF structural and inline elements that were previously created by other *Writers*

Processing Requirements

- XLIFF elements MAY be *Modified* and *Enriched* at the same time.

Note

Extractor or *Enricher* knowledge of the native format is not assumed while *Modifying*.

Modifier, Modifier Agent

an *Agent* that performs the *Modification* process

Translation, Translate

a rendering of the meaning of the source text, expressed in the target language

Writer, Writer Agent

an *Agent* that creates, generates, or otherwise writes an *XLIFF Document* for whatever purpose, including but not limited to *Extractor*, *Modifier*, and *Enricher Agents*.

Note

[ISO 21720:2024](https://standards.iteh.ai/catalog/standards/iso/51b1b925-05b2-4822-9658-6246a29c13dd/iso-21720-2024)

<https://standards.iteh.ai/catalog/standards/iso/51b1b925-05b2-4822-9658-6246a29c13dd/iso-21720-2024>

Since XLIFF is intended as an exchange format rather than a processing format, many applications will need to generate *XLIFF Documents* from their internal processing formats, even in cases when they are processing *XLIFF Documents* created by another *Extractor*.

1.1.3 Key concepts

XLIFF Core

The core of XLIFF 2.1 consists of the minimum set of XML elements and attributes required to (a) prepare a document that contains text extracted from one or more files for localization, (b) allow it to be completed with the translation of the extracted text, and (c) allow the generation of *Translated* versions of the original document.

The XML namespace that corresponds to the core subset of XLIFF 2.1 is "urn:oasis:names:tc:xliff:document:2.0".

XLIFF-defined (elements and attributes)

The following is the list of allowed schema URI prefixes for *XLIFF-defined* elements and attributes:

```
urn:oasis:names:tc:xliff:  
http://www.w3.org/2005/11/its
```

However, the following namespaces are NOT considered *XLIFF-defined* for the purposes of the XLIFF 2.1 specification:

```
urn:oasis:names:tc:xliff:document:1.0  
urn:oasis:names:tc:xliff:document:1.1  
urn:oasis:names:tc:xliff:document:1.2  
urn:oasis:names:tc:xliff:changetracking:2.0
```

Elements and attributes from other namespaces are not *XLIFF-defined*.

XLIFF Document

Any XML document that declares the namespace "urn:oasis:names:tc:xliff:document:2.0" as its main namespace, has `<xliff>` as the root element and complies with the XML Schemas and the declared Constraints that are part of this specification.

XLIFF Module

A module is an OPTIONAL set of XML elements and attributes that stores information about a process applied to an *XLIFF Document* and the data incorporated into the document as result of that process.

Each official module defined for XLIFF 2.1 has its grammar defined in an independent XML Schema with a separate namespace.

1.2 Normative References

[BCP 47] M. Davis, *Tags for Identifying Languages*, <http://tools.ietf.org/html/bcp47> IETF (Internet Engineering Task Force).

[HTML5] Ian Hickos, Robin Berjon, Steve Faulkner, Travis Leithead, Erika Doyle Navara, Edward O'Connor, Silvia Pfeiffer *HTML5. A vocabulary and associated APIs for HTML and XHTML*, <http://www.w3.org/TR/html5/> W3C Recommendation 28 October 2014.

[ITS] David Filip, Shaun McCance, Dave Lewis, Christian Lieske, Arle Lommel, Jirka Kosek, Felix Sasaki, Yves Savourel *Internationalization Tag Set (ITS) Version 2.0*, <http://www.w3.org/TR/its20/> W3C Recommendation 29 October 2013.

[NOTE-datetime] M. Wolf, C. Wicksteed, *Date and Time Formats*, <http://www.w3.org/TR/NOTE-datetime> W3C Note, 15th September 1997.

[NVDL] International Standards Organization, *ISO/IEC 19757-4, Information Technology - Document Schema Definition Languages (DSDL) - Part 4: Namespace-based Validation Dispatching Language (NVDL)*, [http://standards.iso.org/ittf/PubliclyAvailableStandards/c038615_ISO_IEC_19757-4_2006\(E\).zip](http://standards.iso.org/ittf/PubliclyAvailableStandards/c038615_ISO_IEC_19757-4_2006(E).zip) ISO, June 1, 2006.

[RFC 2119] S. Bradner, *Key words for use in RFCs to Indicate Requirement Levels*, <https://www.ietf.org/rfc/rfc2119.txt> IETF (Internet Engineering Task Force) RFC 2119, March 1997.

[RFC 3987] M. Duerst and M. Suignard, *Internationalized Resource Identifiers (IRIs)*, <https://www.ietf.org/rfc/rfc3987.txt> IETF (Internet Engineering Task Force) RFC 3987, January 2005.

[RFC 7303] H. Thompson and C. Lilley, *XML Media Types*, <https://www.tools.ietf.org/html/rfc7303> IETF (Internet Engineering Task Force) RFC 7303, July 2014.

[Schematron] International Standards Organization, *ISO/IEC 19757-3, Information Technology - Document Schema Definition Languages (DSDL) - Part 3: Rule-Based Validation — Schematron (Second Edition)*, http://standards.iso.org/ittf/PubliclyAvailableStandards/c055982_ISO_IEC_19757-3_2016.zip ISO, January 15, 2016.

[UAX #9] M. Davis, A. Lanin, A. Glass, *UNICODE BIDIRECTIONAL ALGORITHM*, <http://www.unicode.org/reports/tr9/tr9-35.html> Unicode Bidirectional Algorithm, May 18, 2016.

[UAX #15] M. Davis, K. Whistler, *UNICODE NORMALIZATION FORMS*, <http://www.unicode.org/reports/tr15/tr15-44.html> Unicode Normalization Forms, February 24, 2016.

[Unicode] The Unicode Consortium, *The Unicode Standard*, <http://www.unicode.org/versions/Unicode9.0.0/> Mountain View, CA: The Unicode Consortium, June 21, 2016.

[XML] W3C, *Extensible Markup Language (XML) 1.0*, <http://www.w3.org/TR/xml/> (Fifth Edition) W3C Recommendation 26 November 2008.

[XML namespace] W3C, *Schema document for namespace* <http://www.w3.org/XML/1998/namespace> <http://www.w3.org/2001/xml.xsd> [<http://www.w3.org/2009/01/xml.xsd>]. at <http://docs.oasis-open.org/xliff/xliff-core/v2.1/os/schemas/informativeCopiesOf3rdPartySchemas/w3c/xml.xsd> in this distribution

[XML Catalogs] Norman Walsh, *XML Catalogs*, <https://www.oasis-open.org/committees/download.php/14809/xml-catalogs.html> OASIS Standard V1.1, 07 October 2005.

[XML Schema] W3C, *XML Schema, refers to the two part standard comprising [XML Schema Structures] and [XML Schema Datatypes]* (Second Editions) W3C Recommendations 28 October 2004.

[XML Schema Datatypes] W3C, *XML Schema Part 2: Datatypes*, <http://www.w3.org/TR/xmlschema-2/> (Second Edition) W3C Recommendation 28 October 2004.

[XML Schema Structures] W3C, *XML Schema Part 1: Structures*, <https://www.w3.org/TR/xmlschema-1/> (Second Edition) W3C Recommendation 28 October 2004.

1.3 Non-Normative References

[LDML] *Unicode Locale Data Markup Language* <http://unicode.org/reports/tr35/>

[SRX] *Segmentation Rules eXchange* <http://www.unicode.org/uli/pas/srx/>

[UAX #29] M. Davis, *UNICODE TEXT SEGMENTATION*, <http://www.unicode.org/reports/tr29/> Unicode text Segmentation.

[XML I18N BP] *Best Practices for XML Internationalization*, 13 February 2008, <http://www.w3.org/TR/xml-i18n-bp/> W3C Working Group.

2 Conformance

1. *Document Conformance*
 - a. XLIFF is an XML vocabulary, therefore conformant *XLIFF Documents* MUST be well formed and valid [XML] documents.
 - b. Conformant *XLIFF Documents* MUST be valid instances of the official Core XML Schema (http://docs.oasis-open.org/xliff/xliff-core/v2.1/os/schemas/xliff_core_2.0.xsd) that is a part of this multipart Work Product.
 - c. As not all aspects of the XLIFF specification can be expressed in terms of XML Schemas, conformant *XLIFF Documents* MUST also comply with all relevant elements and attributes definitions, normative usage descriptions, and Constraints specified in this specification document.
 - d. *XLIFF Documents* MAY contain custom extensions, as defined in the [Extension Mechanisms](#) section.
2. *Application Conformance*
 - a. *XLIFF Writers* MUST create conformant *XLIFF Documents* to be considered XLIFF compliant.
 - b. *Agents* processing conformant *XLIFF Documents* that contain custom extensions are not REQUIRED to understand and process non-XLIFF elements or attributes. However, conformant applications SHOULD preserve existing custom extensions when processing conformant *XLIFF Documents*, provided that the elements that contain custom extensions are not removed according to XLIFF Processing Requirements or the extension's own processing requirements.
 - c. All *Agents* MUST comply with Processing Requirements for otherwise unspecified *Agents* or without a specifically set target *Agent*.
 - d. Specialized *Agents* defined in this specification - this is *Extractor*, *Merger*, *Writer*, *Modifier*, and *Enricher Agents* - MUST comply with the Processing Requirements targeting their specifically defined type of *Agent* on top of Processing Requirements targeting all *Agents* as per point c. above.