TBX Mapping with XLIFF v2.0 and Higher Version 1.0

Working Draft 01

11 April 2017

Specification URIs

This version:
http://docs.oasis-open.org/xliff-omos/xliff-tbx/v1.0/wd01/xliff-tbx-v1.0-wd01.html (Authoritative)
http://docs.oasis-open.org/xliff-omos/xliff-tbx/v1.0/wd01/xliff-tbx-v1.0-wd01.pdf
http://docs.oasis-open.org/xliff-omos/xliff-tbx/v1.0/wd01/xliff-tbx-v1.0-wd01.xml

Previous version:
N/A

Latest version:
http://docs.oasis-open.org/xliff-omos/xliff-tbx/v1.0/xliff-tbx-v1.0.html (Authoritative)
http://docs.oasis-open.org/xliff-omos/xliff-tbx/v1.0/xliff-tbx-v1.0.pdf
http://docs.oasis-open.org/xliff-omos/xliff-tbx/v1.0/xliff-tbx-v1.0.xml

Technical Committee:
OASIS XLIFF Object Model and Other Serializations (XLIFF OMOS) TC

Chair:
David Filip (david.filip@adaptcentre.ie), Trinity College Dublin (ADAPT)

Editors:
David Filip (david.filip@adaptcentre.ie), Trinity College Dublin (ADAPT)
James Hayes (james.s.hayes@byu.edu), Brigham Young University

Additional artifacts:
This prose specification is one component of a Work Product that also includes:

- XML schemas accessible from http://docs.oasis-open.org/xliff-omos/xliff-tbx/v1.0/wd01/schemas/

Related Work:
This specification is related to:


- [related specifications TBX]
Declared XML Namespaces:

- urn:oasis:names:tc:xliff:document:2.0
- urn:oasis:names:tc:xliff:glossary:2.0
- urn:oasis:names:tc:xliff:metadata:2.0

Abstract:
This document defines version 1.0 of TBX Mapping with XLIFF v2.0 and Higher. The purpose of this normative mapping is to ensure that Terminology payload and metadata can be processed and inserted during an XLIFF 2 facilitated roundtrip in way that is semantically aligned with TBX. As result TBX basic Terminology data and metadata can be Extracted into XLIFF Documents, XLIFF Documents can be Enriched by said data and metadata. The processed or newly created Terminology data and metadata can be Merged back again after the XLIFF facilitated roundtrip into the TBX Basic format. The defined mapping is bidirectional, lossless as defined and extensible.

Status:
This document was last revised or approved by the OASIS XLIFF Object Model and Other Serializations (XLIFF OMOS) TC on the above date. The level of approval is also listed above. Check the “Latest version” location noted above for possible later revisions of this document.

Technical Committee members should send comments on this specification to the Technical Committee's email list. Others should send comments to the Technical Committee by using the "Send A Comment" button on the Technical Committee's web page at http://www.oasis-open.org/committees/xliff-omos/.

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 (http://www.oasis-open.org/committees/xliff/ipr.php).

Note for any machine-readable content (aka Computer Language Definitions) declared Normative for this Work Product that is provided in separate plain text files, in the event of a discrepancy between any such plain text file and display content in the Work Product's prose narrative document(s), the content in the separate plain text file prevails.

Citation format:
When referencing this specification the following citation format should be used:

[TBX-XLIFF-Mapping-1.0]
**Table of Contents**

1 Introduction .............................................................................................................................. 5
   1.1 Terminology ................................................................................................................... 5
       1.1.1 Key words ..........................................................................................................  5
       1.1.2 Definitions .......................................................................................................... 5
       1.1.3 Key concepts ...................................................................................................... 7
   1.2 Normative References ..................................................................................................  7
   1.3 Non-Normative References .........................................................................................  8
2 Mapping Specification .............................................................................................................. 9
   2.1 Business Process Description - Why you care for the mapping .................................. 9
   2.2 The Technical Detail - or nuts and bolts ..................................................................  9
       2.2.1 terminological entry ........................................................................................... 9
       2.2.2 term [source] ...................................................................................................... 9
       2.2.3 term [target] ...................................................................................................... 10
       2.2.4 definition .......................................................................................................... 11
       2.2.5 context ............................................................................................................  12
       2.2.6 source .............................................................................................................  12
3 Conformance .......................................................................................................................... 14

**Appendixes**

A Machine Readable Validation Artifacts (Informative) ................................................................. 15
   A.1 XML Schemas Tree .................................................................................................... 15
   A.2 Support Schemas ........................................................................................................ 15
B Specification Change Tracking (Informative) ............................................................................ 17
C Acknowledgements (Informative) ............................................................................................ 18
1 Introduction

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 but also any inline examples or illustrative exemplary lists in otherwise normative text),
Schema and other artifacts listings (the corresponding artifacts are normative, not their listings).

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

Note

The below definitions are restated from XLIFF 2 and higher for reading convenience.

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**

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.

XLIFF 2, XLIFF
In this specification, any mention of XLIFF or XLIFF 2 (unless specified otherwise) means the current OASIS Standard XLIFF Version 2.0 or higher. i.e. the currently OASIS approved and published XLIFF 2 Standard.

**Note**

At the time of development of this specification, XLIFF Version 2.0 was the current OASIS Standard. XLIFF Version 2.1 was entering the second public review. It is not expected that the XLIFF Core or XLIFF Glossary Module would become backwards incompatible in any of the minor feature or bugfix releases of XLIFF 2.
1.1.3 Key concepts

XLIFF Core
The core of XLIFF 2 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 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:
https://www.w3.org/2005/11/its/

However, the following namespaces are NOT considered XLIFF-defined for the purposes of the TBX Mapping with XLIFF v2.0 and Higher Version 1.0 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> [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 1.0 has its grammar defined in an independent XML Schema with a separate namespace.

1.2 Normative References


1.3 Non-Normative References

[XLIFF 2.0 and TBX-Basic] Interoperability of XLIFF 2.0 Glossary Module and TBX-Basic

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

[SRX] Segmentation Rules eXchange
http://www.gala-global.org/oscarStandards/srx/

[UAX #29] M. Davis, UNICODE TEXT SEGMENTATION,

http://www.w3.org/TR/xml-I18n-bp/ W3C Working Group.
2 Mapping Specification

2.1 Business Process Description - Why you care for the mapping

The purpose of the XLIFF Glossary Module is to allow lossless storage of basic glossary information relating to the bi-text contents of an XLIFF file. So as not to surpass the scope and purpose of XLIFF as a whole, this purpose does not extend to serving as a replacement for a more robust terminology management format. In order to better facilitate more detailed interaction with the contents of the Glossary Module, these contents can be converted into a more robust terminological exchange format for use in dedicated termbase management systems. Likewise, it is beneficial to be able to convert basic terminological data from a more detailed terminological exchange format to be used in the Glossary Module, where the data can be linked directly with the text to which it directly applies. With these benefits and the purpose of XLIFF Glossary Module in mind, this mapping has been created to facilitate valid conversion to or from XLIFF Glossary Module and the Basic dialect of TermBase eXchange (TBX-Basic).

2.2 The Technical Detail - or nuts and bolts

2.2.1 terminological entry

XLIFF: glossEntry


TBX: conceptEntry


Constraints

• <tbx:conceptEntry> [http://terminorgs.net/downloads/TBX_Basic_Version_3.1.pdf] MUST have an tbx:id attribute with a unique value beginning with a letter to be valid.

Mapping

Processing Requirements


2.2.2 term [source]

XLIFF: term

TBX: term


Constraints


Mapping

Processing Requirements


2.2.3 term [target]

XLIFF: translation


TBX: term


Constraints


Mapping

Processing Requirements
• Writers MUST convert value such that each XLIFF Glossary Module 
  <gls:translation> [http://docs.oasis-open.org/xliff/xliff-core/v2.0/os/
xliff-core-v2.0-os.html#translation] corresponds to a <tbx:term> [http:/
terminorgs.net/downloads/TBX_Basic_Version_3.1.pdf] descended of the 
  the appropriate language.

• Writers MUST use the value of the <xlf:trgLang> [http://docs.oasis-open.org/
xliff/xliff-core/v2.0/os/xliff-core-v2.0-os.html#trglang] attribute of the 
  <xlf:xliff> [http://docs.oasis-open.org/xliff/xliff-core/v2.0/
os/xliff-core-v2.0-os.html#xliff] element as the value of the 

2.2.4 definition

XLIFF: definition

<gls:definition> [http://terminorgs.net/downloads/TBX_Basic_Version_3.1.pdfw] 
contains definition for the sibling <gls:term> [http://docs.oasis-open.org/xliff/xliff-
core/v2.0/os/xliff-core-v2.0-os.html#term] . See XLIFF Specification [http://docs.oasis-
open.org/xliff/xliff-core/v2.0/os/xliff-core-v2.0-os.html#ref] for more details.

TBX: definition

Definition [http://terminorgs.net/downloads/TBX_Basic_Version_3.1.pdf] is represented as 
<tbx:descrip> [http://terminorgs.net/downloads/TBX_Basic_Version_3.1.pdf] and 
contains a definition for the <tbx:term> [http://terminorgs.net/downloads/TBX_Basic_Version_3.1.pdf] element or elements to which it applies. If <tbx:definition> 
[http://terminorgs.net/downloads/TBX_Basic_Version_3.1.pdf] is placed at the 
level, it applies to all descendant <tbx:term> [http://terminorgs.net/
downloads/TBX_Basic_Version_3.1.pdf] elements of the <tbx:conceptEntry> 
[http://terminorgs.net/downloads/TBX_Basic_Version_3.1.pdf] is placed at the 
<tbx:langSec> [http://terminorgs.net/downloads/TBX_Basic_Version_3.1.pdf] level, 

Constraints

  MUST have a <type> [http://www.tbxinfo.net/wp-content/uploads/2016/10/
tbx_oscar.pdf] attribute with a value of "definition".

• The placement of <tbx:descrip> [http://terminorgs.net/downloads/TBX_Basic_Version_3.1.pdfw] MUST be at either the concept or language levels.

Mapping

Processing Requirements

• Writers MUST convert value such that the <gls:definition> [http://docs.oasis-
open.org/xliff/xliff-core/v2.0/os/xliff-core-v2.0-os.html#definition]
If the `gls:definition` element is accompanied by `gls:source`, Writers MUST ensure that the resulting `<txb:descrip>` is placed in a `<txb:descripGrp>` alongside the mapped value of `gls:source`.

### 2.2.5 context

**XLIFF: ref**

`gls:ref` points to a span of text to which the `gls:glossEntry` is relevant.

**TBX: context**

Context is represented as `<txb:descrip>` and contains authentic context sentences. See TBX Specification for more details.

**Constraints**

- `<txb:descrip>` MUST contain a `txb:type` attribute with a value of "context".
- `<txb:descrip>` MUST be placed at the term level.

**Mapping**

**Processing Requirements**

- Writers MUST convert value such that the `gls:ref` attribute corresponds to a `<txb:descrip>` element.
- Extractors MUST ensure that the full text value of the parent element of the `gls:mrk` element to which the `gls:ref` attribute refers, including the text content of said `gls:mrk` element excluding `gls:mrk` tags, is to be used as the text content of the `<txb:descrip>` element.

### 2.2.6 source

**XLIFF: source**
<gls:source> [http://docs.oasis-open.org/xliff/xliff-core/v2.0/os/xliff-core-v2.0-os.html#gls_source] is an attribute which can be added to an element and indicates the origin of the element to which it is applied. See XLIFF Specification [http://docs.oasis-open.org/xliff/xliff-core/v2.0/os/xliff-core-v2.0-os.html#gls_source] for more details.

**TBX: source**


**Constraints**


**Mapping**

**Processing Requirements**


3 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/wd01/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 [extensions.xml] 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.
   e. XLIFF is a format explicitly designed for exchanging data among various Agents. Thus, a conformant XLIFF application MUST be able to accept XLIFF Documents it had written after those XLIFF Documents were Modified or Enriched by a different application, provided that:
      i. The processed files are conformant XLIFF Documents,
      ii. in a state compliant with all relevant Processing Requirements.

3. Backwards Compatibility
   a. Conformant applications are REQUIRED to support XLIFF 2.0.
   b. Conformant applications are NOT REQUIRED to support XLIFF 1.2 or previous Versions.

Note

XLIFF Documents conformant to this specification are not and cannot be conformant to XLIFF 1.2 or earlier versions. If an application needs to support for whatever business reason both XLIFF 2 and XLIFF 1.2 or earlier, these will need to be supported as separate functionalities.
Appendix A Machine Readable Validation Artifacts (Informative)

This appendix summarizes information on machine readable validation artifacts for XLIFF Core and its Glossary Module as well as TBX Basic [Module and TBX Basic artifacts Info needs added]

1. XLIFF Core XML Schema,
2. XML Catalogue of XLIFF Defined XML Schemas,
3. Other XLIFF Defined validation artcifacts such as Schematron Schemas, NVDL schema (relax NG schemas?) etc.

The basic grammar of XLIFF 1.0 is defined using [actual number of schemas - nine (9)]XML Schemas and one (1) XML catalog. The module schemas are referenced from their respective modules.

Advanced static Constraints and dynamic Processing Requirements that could not be expressed using XML Schema 1.0 [add normative reference] are expressed in Schematron [add normative reference] schemas.

Relationships among all of the above mentioned machine redable validation artifacts provided as part of this multipart product is expressed using one (1) NVDL [add normative reference] schema.

[Relax NG schemas of XLIFF Core and Modules are provided. Although the Relax NG schemas are intended as normative validation artifacts, the XML Schema 1.0 [add normative reference] schemas and Schematron [add normative reference] schemas have precedence in case of conflicting validation results.] [This is only relevant if Relax NG is provided.]

A.1 XML Schemas Tree

```
Core XML Schema [http://docs.oasis-open.org/xliff/xliff-core/v2.1/wd01/schemas/xliff_core_2.1.xsd]
  |---Candidates Module XML Schema [http://docs.oasis-open.org/xliff/xliff-core/v2.1/wd01/schemas/modules/matches.xsd]
  |---Glossary Module XML Schema [http://docs.oasis-open.org/xliff/xliff-core/v2.1/wd01/schemas/modules/glossary.xsd]
  |---Format Style Module XML Schema [http://docs.oasis-open.org/xliff/xliff-core/v2.1/wd01/schemas/modules/fs.xsd]
  |---Metadata Module XML Schema [http://docs.oasis-open.org/xliff/xliff-core/v2.1/wd01/schemas/modules/metadata.xsd]
  |---Change Tracking Module XML Schema [http://docs.oasis-open.org/xliff/xliff-core/v2.1/wd01/schemas/modules/change_tracking.xsd]
  |---Size and Length Restriction Module XML Schema [http://docs.oasis-open.org/xliff/xliff-core/v2.1/wd01/schemas/modules/size_restriction.xsd]
  |---Validation Module XML Schema [http://docs.oasis-open.org/xliff/xliff-core/v2.1/wd01/schemas/modules/validation.xsd]
```

A.2 Support Schemas

Third party support schemas that are normatively referenced from this specification or from the machine readable artifacts that are a part of this multipart product are distributed along with the XLIFF-defined schemas in a subfolder named informativeCopiesOf3rdPartySchemas and further subdivided in folders according to the owner/maintainer of the schema.
Warning

Schema copies in this sub-folder are provided solely for implementers convenience and are NOT a part of the OASIS multipart product. These schemas belong to their respective owners and their use is governed by their owners’ respective IPR policies. The support schemas are organized in folders per owner/maintainer. It is the implementer's sole responsibility to ensure that their local copies of all schemas are the appropriate up to date versions.

Currently the only included third party support schema is http://www.w3.org/2001/xml.xsd [http://www.w3.org/2009/01/xml.xsd] at http://docs.oasis-open.org/xliff-omos/xliff-tbx/v1.0/wd01/schemas/informativeCopiesOf3rdPartySchemas/w3c/xml.xsd in this distribution.
Appendix B Specification Change Tracking (Informative)

This appendix will contain tracked changes after the csprd01 phase will have been reached.
Appendix C Acknowledgements (Informative)

The following individuals have participated in the creation of this specification and are gratefully acknowledged:

- Filip, David - TCD, ADAPT Centre
- Hayes, James - Brigham Young University
- King, Ryan - Microsoft
- Morado Vázquez, Lucía - University of Geneva
- Phil Ritchie - Vistatec
- Soroush Saadatfar, Localisation Research Centre
- Felix Sasaki - Individual
- Savourel, Yves - ENLASO Corporation
- Schnabel, Bryan - Individual
- Tingley, Chase - Spartan Software Inc.