**XLIFF 2.0 Mapping**

Jump to: [navigation](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#mw-navigation), [search](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#p-search)

**Contents**

* [1 Introduction](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Introduction)
* [2 Testing the mapping](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Testing_the_mapping)
* [3 Data Categories Existing in XLIFF](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Data_Categories_Existing_in_XLIFF)
	+ [3.1 Translate](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Translate)
		- [3.1.1 Structural Elements](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Structural_Elements)
		- [3.1.2 For inline elements](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#For_inline_elements)
	+ [3.2 Preserve Space (==========TO REVIEW)](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Preserve_Space_.28.3D.3D.3D.3D.3D.3D.3D.3D.3D.3DTO_REVIEW.29)
		- [3.2.1 Structural Elements](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Structural_Elements_2)
		- [3.2.2 Inline Elements](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Inline_Elements)
* [4 Data Categories Partially Covered in XLIFF](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Data_Categories_Partially_Covered_in_XLIFF)
	+ [4.1 Localization Note (==========TODO)](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Localization_Note_.28.3D.3D.3D.3D.3D.3D.3D.3D.3D.3DTODO.29)
		- [4.1.1 Structural Elements](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Structural_Elements_3)
		- [4.1.2 Inline Elements](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Inline_Elements_2)
	+ [4.2 Terminology (==========TO REVIEW)](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Terminology_.28.3D.3D.3D.3D.3D.3D.3D.3D.3D.3DTO_REVIEW.29)
		- [4.2.1 Structural Elements](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Structural_Elements_4)
		- [4.2.2 Inline Elements](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Inline_Elements_3)
	+ [4.3 Language Information (==========TO REVIEW)](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Language_Information_.28.3D.3D.3D.3D.3D.3D.3D.3D.3D.3DTO_REVIEW.29)
		- [4.3.1 Structural Elements](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Structural_Elements_5)
		- [4.3.2 Inline Elements](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Inline_Elements_4)
	+ [4.4 MT Confidence (==========TO REVIEW)](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#MT_Confidence_.28.3D.3D.3D.3D.3D.3D.3D.3D.3D.3DTO_REVIEW.29)
		- [4.4.1 Structural Elements](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Structural_Elements_6)
			* [4.4.1.1 In the match element](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#In_the_match_element)
		- [4.4.2 Inline Elements](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Inline_Elements_5)
* [5 Data Categories Represented Using ITS Itself](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Data_Categories_Represented_Using_ITS_Itself)
	+ [5.1 Domain (==========TO REVIEW)](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Domain_.28.3D.3D.3D.3D.3D.3D.3D.3D.3D.3DTO_REVIEW.29)
		- [5.1.1 Structural Elements](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Structural_Elements_7)
		- [5.1.2 Inline Elements](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Inline_Elements_6)
	+ [5.2 Text Analysis (==========TO REVIEW)](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Text_Analysis_.28.3D.3D.3D.3D.3D.3D.3D.3D.3D.3DTO_REVIEW.29)
		- [5.2.1 Structural Elements](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Structural_Elements_8)
		- [5.2.2 Inline Elements](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Inline_Elements_7)
	+ [5.3 Locale Filter (==========TO REVIEW)](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Locale_Filter_.28.3D.3D.3D.3D.3D.3D.3D.3D.3D.3DTO_REVIEW.29)
		- [5.3.1 Structural Elements](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Structural_Elements_9)
			* [5.3.1.1 When the Target Locale in XLIFF is Undefined](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#When_the_Target_Locale_in_XLIFF_is_Undefined)
			* [5.3.1.2 When the Target Locale in XLIFF is Defined](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#When_the_Target_Locale_in_XLIFF_is_Defined)
		- [5.3.2 Inline Elements](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Inline_Elements_8)
			* [5.3.2.1 When the Target Locale in XLIFF is Undefined](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#When_the_Target_Locale_in_XLIFF_is_Undefined_2)
			* [5.3.2.2 When the Target Locale in XLIFF is Defined](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#When_the_Target_Locale_in_XLIFF_is_Defined_2)
	+ [5.4 Provenance (==========TO REVIEW)](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Provenance_.28.3D.3D.3D.3D.3D.3D.3D.3D.3D.3DTO_REVIEW.29)
		- [5.4.1 Structural Elements](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Structural_Elements_10)
		- [5.4.2 Inline Elements](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Inline_Elements_9)
	+ [5.5 Localization Quality Issue (==========TO REVIEW)](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Localization_Quality_Issue_.28.3D.3D.3D.3D.3D.3D.3D.3D.3D.3DTO_REVIEW.29)
		- [5.5.1 Structural Elements](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Structural_Elements_11)
		- [5.5.2 Inline Elements](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Inline_Elements_10)
	+ [5.6 Allowed Characters (==========TO REVIEW)](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Allowed_Characters_.28.3D.3D.3D.3D.3D.3D.3D.3D.3D.3DTO_REVIEW.29)
		- [5.6.1 Structural Elements](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Structural_Elements_12)
		- [5.6.2 Inline Elements](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Inline_Elements_11)
* [6 Data Categories Not Representing Metadata](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Data_Categories_Not_Representing_Metadata)
	+ [6.1 Element Within Text (==========TO REVIEW)](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Element_Within_Text_.28.3D.3D.3D.3D.3D.3D.3D.3D.3D.3DTO_REVIEW.29)
	+ [6.2 Target Pointer (==========TO REVIEW)](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Target_Pointer_.28.3D.3D.3D.3D.3D.3D.3D.3D.3D.3DTO_REVIEW.29)
	+ [6.3 Id Value (==========TO REVIEW)](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Id_Value_.28.3D.3D.3D.3D.3D.3D.3D.3D.3D.3DTO_REVIEW.29)
		- [6.3.1 Structural Elements](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Structural_Elements_13)
		- [6.3.2 Inline Elements](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Inline_Elements_12)
* [7 Data Categories Not Mapped Yet](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Data_Categories_Not_Mapped_Yet)
	+ [7.1 Directionality (==========TODO)](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Directionality_.28.3D.3D.3D.3D.3D.3D.3D.3D.3D.3DTODO.29)
		- [7.1.1 Structural Elements](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Structural_Elements_14)
		- [7.1.2 Inline Elements](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Inline_Elements_13)
	+ [7.2 External Resource (==========TODO)](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#External_Resource_.28.3D.3D.3D.3D.3D.3D.3D.3D.3D.3DTODO.29)
		- [7.2.1 Structural Elements](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Structural_Elements_15)
		- [7.2.2 Inline Elements](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Inline_Elements_14)
	+ [7.3 Localization Quality Rating (==========TODO)](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Localization_Quality_Rating_.28.3D.3D.3D.3D.3D.3D.3D.3D.3D.3DTODO.29)
		- [7.3.1 Structural Elements](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Structural_Elements_16)
		- [7.3.2 Inline Elements](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Inline_Elements_15)
	+ [7.4 Storage Size (==========TODO)](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Storage_Size_.28.3D.3D.3D.3D.3D.3D.3D.3D.3D.3DTODO.29)
		- [7.4.1 Structural Elements](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Structural_Elements_17)
		- [7.4.2 Inline Elements](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Inline_Elements_16)
* [8 References](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#References)

**Introduction**

This document provides a recommendation on how the ITS 2.0 data categories are represented in XLIFF 2.
For the mapping between ITS 2.0 and XLIFF 1.2 see the page "[XLIFF 1.2 Mapping](http://www.w3.org/International/its/wiki/XLIFF_1.2_Mapping)".

Notes:

* Please, use the IG mailing list (<http://lists.w3.org/Archives/Public/public-i18n-its-ig/>) for discussing this topic.
* The 'structural' entries relate to the cases where the element with the ITS information is a non-inline (structural) element. For example a <p> in HTML.
* The 'inline' entries relate to the case where the element with the ITS information is an inline element. For example a <span> in HTML.
* The prefix itsxlf refers to the namespace <http://www.w3.org/ns/its-xliff/>

ITS data categories can be classified into several categories in XLIFF:

* [Data Categories Existing in XLIFF](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Data_Categories_Existing_in_XLIFF)
* [Data Categories Partially Covered in XLIFF](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Data_Categories_Partially_Covered_in_XLIFF)
* [Data Categories Represented Using ITS Itself](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Data_Categories_Represented_Using_ITS_Itself)
* [Data Categories Not Representing Metadata](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Data_Categories_Not_Representing_Metadata)
* And because this document is still under definition there are: [Data Categories Not Mapped Yet](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping#Data_Categories_Not_Mapped_Yet)

**Testing the mapping**

This section is a stub. Feel free to complete it by providing e.g. these ideas:

* What input files are needed: XLIFF, general XML, HTML5?
	+ XLIFF 2 documents, and I suppose (to see if an extractor supports ITS too): HTML5 or XML documents
* What output is needed: XLIFF only?
	+ For the extraction case: the XLIFF output
	+ But I suppose some kind of comparable text format would be ideal. I'm not sure the same XPath-based format we used for ITS would be best here as XLIFF processors may be using very different way to process the document. Maybe something using the ID of the object rather than the path would be better?
* How would the conformance of the output to mappings be tested?
	+ Ideally by comparing the gold output to the tool's output
* What would be a good location of the test files - a github repository or a XLIFF / ITS group specific location ? Advantage: many people can contribute
	+ Github would be fine. I'm guessing there may be some call for hosting this in SVN's OASIS too.

**Data Categories Existing in XLIFF**

**Translate**

Indicates whether a content is translatable or not.
See <http://www.w3.org/TR/its20/#trans-datacat> for details.

**Structural Elements**

Use the translate attribute:

Original:

<p translate='yes|no'>Text</p>

Extraction:

<unit id='1' translate='yes|no'>

 <segment>

 <source>Text</source>

 </segment>

</unit>

If the element is not translatable you can also simply not extract it.

**For inline elements**

Use <mrk> with translate='yes|no'>. A fall-back option is to extract the non-translatable content as inline code,

Original:

<p>Text <code translate='no'>Code</code></p>

Extraction:

<unit id='1'>

 <segment>

 <source>Text <pc id='1'/><mrk id='m1' translate='no'>Code</mrk></pc></source>

 </segment>

</unit>

or

<unit id='1'>

 <segment>

 <source>Text <ph id='1'/></source>

 </segment>

</unit>

**Preserve Space (==========TO REVIEW)**

Indicates how whitespace should be handled in a given content.
See <http://www.w3.org/TR/its20/#preservespace> for more details.

**Structural Elements**

Whitespace handling at the structural level is indicated with xml:space in XLIFF 2:

Original:

<listing xml:space='preserve'>Line 1

Line 2</listing>

Extraction:

<unit id='1' xml:space='preserve'>

 <segment>

 <source>Line 1

Line 2</source>

 </segment>

</unit>

**Inline Elements**

Use the attribute xml:space in <mrk>.

Original:

<para>Normal text and

 <span xml:space="preserve">preserved spaces: [ ]</span>.

</para>

Extraction:

<unit id='1'>

 <segment>

 <source>Normal text and <pc id='1'><mrk

 xml:space="preserve" mtype='its:any'>preserved spaces: [ ]</mrk></pc>.</source>

 </segment>

</unit>

Note that, currently, few localization applications will honor preserving whitespace for only a given span of text.

**Data Categories Partially Covered in XLIFF**

**Localization Note (==========TODO)**

Provides a way to communicate notes to localizers about a particular item of content.
See <http://www.w3.org/TR/its20/#locNote-datacat> for more details.

**Structural Elements**

TODO

**Inline Elements**

TODO

**Terminology (==========TO REVIEW)**

Marks terms and optionally associates them with information, such as definitions.
See <http://www.w3.org/TR/its20/#terminology> for more details.

**Structural Elements**

It is recommended to map terminology information that appears on a structural element in the original document by using an inline <mrk> element.

Original:

<p its-term='yes'>Term</p>

Extracted:

<unit id='1'>

 <segment>

 <source><mrk id="m1" type='term'>Term</mrk></source>

 </segment>

</unit>

**Inline Elements**

In XLIFF 2 terms are denoted using <mrk type='term'>:

Original:

<p>Text with a <span its-term='yes'>term</span>.</p>

Extracted:

<unit id='1'>

 <segment>

 <source>Text with a <pc id='1'><mrk id="m1" type='term'>term</mrk></pc>.</source>

 </segment>

</trans-unit>

* Use type="its:term-no" for denoting instances where you have its:term="no".
* its:termInfoRef is mapped to the XLIFF ref attribute.
* its:termConfidence is mapped to itsxlf:termConfidence.
* When itsxlf:termConfidence is used, the annotated text MUST be contained within an element with a relevant its:annotatorsRef.
* The attribute value can be used to store information denoted by the global rule attribute its:termInfoPointer.

**WARNING: TBD: the XLIFF 2 specification allow ref and value to be both set at the same time. ITS 2.0 does not allow an info and an info-ref to be set at the same time. So we have to decide something for this case.**

Note: If needed, the value of the ITS termInfoRef attribute is to be adjusted to point to a resource accessible from the XLIFF document. The location and format of this resource is decided by the tool creating the XLIFF document.

Original:

<p>Text with a <span its-term='yes' its-term-info-ref='http://en.wikipedia.org/wiki/Terminology'

 its-term-confidence='0.9'>term</span>.</p>

Extracted:

<unit id='1' its:annotatorsRef='terminology|http://www.cngl.ie/termchecker'>

 <segment>

 <source>Text with a <pc id='1'><mrk id='m1' type='term'

 itsxlf:termInfoRef='http://en.wikipedia.org/wiki/Terminology'

 itsxlf:termConfidence='0.9'>term</mrk></pc>.</source>

 </segment>

</unit>

**Language Information (==========TO REVIEW)**

Expresses the language for a given content.
See <http://www.w3.org/TR/its20/#language-information> for more details.

XLIFF is a bilingual document and defines the source and target language of its payload using the srcLang and trgLang attributes in the <xliff> element. By default, those languages apply to the content of <source> and <target>.

**Structural Elements**

Because XLIFF documents are normally source-monolingual, whole paragraphs in the source document that are not in the main source language are generally not to be extracted.

If there is a need to extract such content, the XLIFF output has to use an inline <mrk> element to enclose the content in a different language than the normal source language of the document.

**Inline Elements**

Use the attribute xml:lang in <mrk>.

Original:

<!doctype html>

<html lang="en">

 <head>

 <meta charset="utf-8">

 <title>My Document</title>

 </head>

 <body>

 <p>Span of text <span lang="fr">en français</span>.</p>

 </body>

</html>

Extraction:

...

<unit id='2'>

 <segment>

 <source>Span of text <pc id='1'><mrk id="m1" xml:lang="fr" mtype='its:any'

 >en français</mrk></pc>.</source>

 </segment>

</unit>

...

**MT Confidence (==========TO REVIEW)**

Communicates the self-reported confidence score from a machine translation engine of the accuracy of a translation it has provided.
See <http://www.w3.org/TR/its20/#mtconfidence> for more details.

**Structural Elements**

It is not recommended that MT Confidence be used at a structural level.

If a structural element of the original document has an MT Confidence annotation, it is recommended to represent that annotation using a <mrk> element that encloses the whole content of the <source> element. See the Inline Elements section below for details.

The MT Confidence score must be within the scope of a corresponding its:annotatorsRef attribute.

**In the match element**

The MT Confidence data category can also be used on the <match> element of the Translation Candidates module. In that case: use the matchQuality attribute to store the value. You must adjust the value by multiplying it by 100 as the scale of matchQuality is [0.0 to 100.0] and the scale for the MT Confidence is [0.0 to 1.0].

<xliff xmlns="urn:oasis:names:tc:xliff:document:2.0"

 xmlns:mtc="urn:oasis:names:tc:xliff:matches:2.0"

 xmlns:its="<http://www.w3.org/2005/11/its>" its:version="2.0"

 version="2.0" srcLang="en" trgLang="fr">

<file id="f1" its:annotatorsRef="mt-confidence|MTServices-XYZ">

 <unit id="1">

 <mtc:matches>

 <!-- Score provided by MTServices-XYZ -->

 <mtc:match ref="#m1" matchQuality="89.82">

 <source>Text</source>

 <target >Texte</target>

 </mtc:match>

 <!-- Score provided by MTProvider-ABC -->

 <mtc:match ref="#m1" matchQuality="67.8"

 its:annotatorsRef="mt-confidence|MTProvider-ABC">

 <source>Text</source>

 <target >Texte</target>

 </mtc:match>

 <!-- Score provided by MTProvider-JKL -->

 <mtc:match ref="#m1" matchQuality="65"

 its:annotatorsRef="mt-confidence|MTProvider-JKL">

 <source>Text</source>

 <target >texte</target>

 </mtc:match>

 <!-- Score provided by MTServices-XYZ -->

 <mtc:match ref="#m1" matchQuality="89.82">

 <source>Some text</source>

 <target>Du texte</target>

 </mtc:match>

 </mtc:matches>

 <segment>

 <source><mrk id='m1' type='mtc:match'>Text</mrk></source>

 </segment>

 </unit>

</file>

</xliff>

Note that matchQuality cannot be mapped to ITS MT Confidence directly as no its:mtConfidencePointer is defined in ITS 2.0.

**Inline Elements**

Use the its:mtConfidence attributes on the <mrk> element.

<target>

 <mrk id="m1" type="its:any"

 its:mtConfidence="0.8982"

 its:annotatorsRef="mt-confidence|MTServices-XYZ"

 >Some translated text</mrk>

</target>

**Data Categories Represented Using ITS Itself**

**Domain (==========TO REVIEW)**

Identifies the topic or subject of a given content.
See <http://www.w3.org/TR/its20/#domain> for more details.

**Structural Elements**

Use the attribute itsxlf:domains:

Original:

<!doctype html>

<html lang="en">

 <head>

 <meta charset="utf-8">

 <title>Data Category: Domain</title>

 <script type="application/its+xml">

 <its:rules xmlns:its="<http://www.w3.org/2005/11/its>" version="2.0"

 xmlns:h="<http://www.w3.org/1999/xhtml>">

 <its:domainRule selector="//h:\*[@class='dom1']" domainPointer="./@class"

 domainMapping="dom1 domain1" />

 </its:rules>

 </script>

 </head>

 <body>

 <p class="dom1">Text in the domain domain1</p>

 </body>

</html>

Extraction:

...

<unit id='2' itsxlf:domains='domain1'>

 <segment>

 <source>Text in the domain domain1</source>

 </segment>

</unit>

...

**Inline Elements**

Use the attribute itsxlf:domains in <mrk>:

Original:

<!doctype html>

<html lang="en">

 <head>

 <meta charset="utf-8">

 <title>Data Category: Domain</title>

 <script type="application/its+xml">

 <its:rules xmlns:its="<http://www.w3.org/2005/11/its>" version="2.0"

 xmlns:h="<http://www.w3.org/1999/xhtml>">

 <its:domainRule selector="//h:\*[@class='dom1']" domainPointer="./@class"

 domainMapping="dom1 domain1" />

 </its:rules>

 </script>

 </head>

 <body>

 <p>Span of text <span class="dom1">in the domain domain1</span></p>

 </body>

</html>

Extraction:

...

<unit id='2'>

 <segment>

 <source>Span of text <pc id='1'><mrk id='m1' mtype='its:any' itsxlf:domains='domain1'

 >in the domain domain1</mrk></pc></source>

 </segment>

</trans-unit>

...

**Text Analysis (==========TO REVIEW)**

Annotates content with lexical or conceptual information for the purpose of contextual disambiguation.
See <http://www.w3.org/TR/its20/#textanalysis> for more details.

**Structural Elements**

Text Analysis is not to be used at a structural level.

If a structural element of the original document has a Text Analysis annotation, it is RECOMMENDED to represent that annotation using a <mrk> element that encloses the whole content of the <source> element.

Original:

<p its-ta-class-ref="<http://nerd.eurecom.fr/ontology#Place>"

 its-ta-ident-ref="<http://dbpedia.org/resource/Arizona>">Arizona</p>

Extraction:

<unit id="1">

 <segment>

 <source><mrk id="m1" type="its:any"

 its:taClassRef="<http://nerd.eurecom.fr/ontology#Place>"

 its:taIdentRef="<http://dbpedia.org/resource/Arizona>">Arizona</mrk></source>

 </segment>

</unit>

**Inline Elements**

Use the ITS attributes in the <mrk> element.

If its:taConfidence is used, the annotated text must be contained within an element with a relevant its:annotatorsRef.

Original:

<div its-annotators-ref="text-analysis|<http://enrycher.ijs.si>">

...

 <p><span its-ta-class-ref="<http://nerd.eurecom.fr/ontology#Place>"

 its-ta-ident-ref="<http://dbpedia.org/resource/Arizona>">Arizona</span></p>

...

</div>

Extraction:

<unit id="1" its:annotatorsRef="text-analysis|<http://enrycher.ijs.si>">

 <segment>

 <source><mrk id="m1" type="its:any"

 its:taClassRef="<http://nerd.eurecom.fr/ontology#Place>"

 its:taIdentRef="<http://dbpedia.org/resource/Arizona>">Arizona</mrk></source>

 </segment>

</unit>

**Locale Filter (==========TO REVIEW)**

Specifies that a content is only applicable to certain locales.
See <http://www.w3.org/TR/its20/#LocaleFilter> for more details.

**Structural Elements**

**When the Target Locale in XLIFF is Undefined**

Use ITS attributes:

Original:

<p its-locale-filter-list='fr'>Text A</p>

<p its-locale-filter-list='ja'>Text B</p>

Extraction:

<xliff srcLang='en' ...>

...

<unit id='1' its:localeFilterList='fr'>

 <segment>

 <source>Text A</source>

 </segment>

</unit>

<unit id='2' its:localeFilterList='ja'>

 <segment>

 <source>Text B</source>

 </segment>

</unit>

**When the Target Locale in XLIFF is Defined**

Use the translate attribute (yes if the target locale applies, no if it does not). It is also recommended to keep the original ITS attributes, so the file could potentially be re-purposed (even if it has a current target):

Original:

<p its-locale-filter-list='fr'>Text A</p>

<p its-locale-filter-list='ja'>Text B</p>

Extraction:

<xliff srcLang='en' trgLang='fr' ...>

...

<unit id='1' translate='yes' its:localeFilterList='fr'>

 <segment>

 <source>Text A</source>

 </segment>

</unit>

<unit id='2' translate='no' its:localeFilterList='ja'>

 <segment>

 <source>Text B</source>

 </segment>

</unit>

If the entry does not apply to the defined target locale you can also simply not extract it.

**Inline Elements**

**When the Target Locale in XLIFF is Undefined**

Use the <mrk> element with the original ITS attributes:

Original:

<p>Text <span its-locale-filter-list='fr' its-locale-filter-type='exclude'>text</span></p>

Extraction:

<xliff srcLang='en' ...>

...

<unit id='1'>

 <segment>

 <source>Text <pc id='1'><mrk id='m1' type='its:any'

 its:localeFilterList='fr' its:localeFilterType='exclude'>text</mrk></g></source>

 </segment>

</unit>

**When the Target Locale in XLIFF is Defined**

Use the <mrk> element with translate='yes' if the target does apply or translate='no' if it does not. It is also recommended to keep the original ITS attributes, so the file could potentially be re-purposed (even if it has a current target).

Original:

<p>Text <span its-locale-filter-list='fr' its-locale-filter-type='exclude'>text</span></p>

Extraction:

<xliff srcLang='en' trgLang='fr'...>

...

<unit id='1'>

 <segment>

 <source>Text <pc id='1'><mrk id='m1' type='its:any' translate='no'

 its:localeFilterList='fr' its:localeFilterType='exclude'>text</mrk></g></source>

 </segment>

</unit>

If the content does not apply to the defined target locale you can also simply replace it by an inline code.

**Provenance (==========TO REVIEW)**

Communicates the identity of agents that have been involved in the translation of the content or the revision of the translated content.
See <http://www.w3.org/TR/its20/#provenance> for more details.

**Structural Elements**

The Provenance data category can be used on <file>, <group> and <unit>.

If a standoff element is needed (because the annotated element has more than one set of the provenance attributes), the <its:provenanceRecords> element must be located in same the element as where the reference is declared.

<unit id='1' its:provenanceRecordsRef="#its=prov1">

 <its:provenanceRecords xml:id="prov1">

 <its:provenanceRecord person="John Doe"/>

 <its:provenanceRecord revPerson="John Smith"/>

 </its:provenanceRecords>

...

**Inline Elements**

For annotating the source or the target content, use the <<mrk> element with the ITS attributes. If a standoff <its:provenanceRecords> element is being used, it must be located in the same <unit> as where the inline rference is declared.

<unit id='1'>

 <its:provenanceRecords xml:id="prov1">

 <its:provenanceRecord person="John Doe"/>

 <its:provenanceRecord revPerson="John Smith"/>

 </its:provenanceRecords>

 <segment>

 <source>Some text</source>

 <target><mrk id='m1' type='its:any' its:provenanceRecordsRef="#its=prov1">Some text</mrk></target>

 </segment>

</unit>

**Localization Quality Issue (==========TO REVIEW)**

Expresses information related to localization quality assessment tasks.
See <http://www.w3.org/TR/its20/#lqissue> for more details.

**Structural Elements**

Localization Quality Issue annotation may be used to annotate the source or the target content within a <unit> element. It is done by using the <mrk> element. See below for details.

**Inline Elements**

The ITS attributes for Localization Quality Issue may be used inline with an <mrk> within a <source> or <target> elements in a <unit> element. For example for a single instance of the Localization Quality Issue data category:

<unit id="1">

 <segment>

 <source>This is the content</source>

 <target><mrk id="m1" type="its:any" its:locQualityIssueType="misspelling"

 its:locQualityIssueComment="'c'es' is unknown. Could be 'c'est'"

 its:locQualityIssueSeverity="50">c'es</mrk> le contenu</target>

 </segment>

</unit>

When needed, a stand-off notation can be used and it is located at the unit's extension point (before the first <segment> element). Note that the reference must used the XLIFF's fragment identifier syntax. The Fragment identifier prefix for the ITS module/extension is its.

<unit id="1">

 <its:locQualityIssues xml:id="lqi1">

 <its:locQualityIssue

 locQualityIssueType="misspelling"

 locQualityIssueComment="'c'es' is unknown. Could be 'c'est'"

 locQualityIssueSeverity="50" />

 <its:locQualityIssue

 locQualityIssueType="grammar"

 locQualityIssueComment="Sentence is not capitalized"

 locQualityIssueSeverity="20" />

 </its:locQualityIssues>

 <segment>

 <source>This is the content</source>

 <target><mrk id="m1" type="its:any"

 its:locQualityIssuesRef="#its=lqi1">c'es le contenu</mrk></target>

 </segment>

</unit>

**Allowed Characters (==========TO REVIEW)**

Specifies the characters that are permitted in a given piece of content.
See <http://www.w3.org/TR/its20/#allowedchars> for more details.

**Structural Elements**

If a structural element of the original document has a Allowed Characters annotation, it is recommended to represent that annotation using a <mrk> element that encloses the whole content of the <source> element. For example:

Original:

<p its-allowedCharacters="[a-ZA-Z]">Text</p>

Extraction:

<unit id="1">

 <segment>

 <source><mrk id="m1" type="its:any" its:allowedCharacters="[a-ZA-Z]">Text</source>

 </segment>

</unit>

**Inline Elements**

Use the ITS attribute on the <mrk> element:

<unit id="1">

 <segment>

 <source>user name:

 <pc id="1">&lt:mrk id="m1" type="its:any" its:allowedCharacters="[a-ZA-Z]">johnDoe</mrk></pc>.</source>

 </segment>

</unit>

**Data Categories Not Representing Metadata**

**Element Within Text (==========TO REVIEW)**

Indicates if an element should be treated as part of a text flow, or as a separate "paragraph".
See <http://www.w3.org/TR/its20/#elements-within-text> for more details.

This data category is not used directly in XLIFF, but it drives what XLIFF element is used to represent the original element in the extracted document:

* withinText='no': Use <unit>
* withinText='yes': Use an inline element such as <pc>, <sc>/<ec> or <ph>.
* withinText='nested': Use a separate <unit>.

**Target Pointer (==========TO REVIEW)**

Provides a way to associate the node of a given source content (i.e. the content to be translated) and the node of its corresponding target content.
See <http://www.w3.org/TR/its20/#target-pointer> for more details.

This data category is not mapped to XLIFF but used by extracting and merging tools to get the source content from the original document and put back the translated content at its proper location.

Note that ITS processors working on XLIFF documents should use the following rule to locate the source and target content:

<its:targetPointerRule selector="//xlf:source" targetPointer="../xlf:target"/>

**Id Value (==========TO REVIEW)**

Indicates a value that can be used as unique identifier for a given part of the content.
See <http://www.w3.org/TR/its20/#idvalue> for more details.

Note that the identifiers in XLIFF are not unique per document, so using the Id Value data category to specify IDs in an XLIFF document is largely useless, excepted when used in very specific contexts that cannot be expressed in the ITS rules. See the [Fragment Identifier section](http://docs.oasis-open.org/xliff/xliff-core/v2.0/xliff-core-v2.0.html#fragid) for details on IDs in XLIFF 2.

**Structural Elements**

Use the name attribute in <unit>:

Original:

<p id='p1'>Text of the paragraph.</p>

Extraction:

<unit id='1' name='p1'>

 <segment>

 <source>Text of the paragraph.</source>

 </segment>

</unit>

**Inline Elements**

The Id Value data category is not mapped to inline codes.

**Data Categories Not Mapped Yet**

**Directionality (==========TODO)**

Provides information about the text directionality of the content.
See <http://www.w3.org/TR/its20/#directionality> for more details.

**Structural Elements**

TODO

**Inline Elements**

TODO

**External Resource (==========TODO)**

Indicates that a node represents or references potentially translatable data in a resource outside the document.
See <http://www.w3.org/TR/its20/#externalresource> for more details.

**Structural Elements**

Use the attribute itsxlf:externalResourceRef in <trans-unit>:

Original:

<its:rules version="2.0" xmlns:its="<http://www.w3.org/2005/11/its>"

 xmlns:html="<http://www.w3.org/1999/xhtml>">

 <its:externalResourceRefRule selector="//html:video/@src"

 externalResourceRefPointer="."/>

 <its:externalResourceRefRule selector="//html:video/@poster"

 externalResourceRefPointer="."/>

</its:rules> ..

<video

 height=360

 poster=video-image.png

 src=<http://www.example.com/video/v2.mp>

 width=640>

Extraction:

...

<trans-unit id='2' itsxlf:externalResourceRef="<http://www.example.com/video/v2.mp>">

...

**Inline Elements**

Use the attribute itsxlf:externalResourceRef in the inline element that holds the reference (e.g. <x/> or <ph>):

Original:

<!doctype html>

<html lang="en">

 <head>

 <meta charset="utf-8">

 <title>Data Category: External Resource</title>

 <script type="application/its+xml">

 <its:rules xmlns:its="<http://www.w3.org/2005/11/its>" version="2.0" xmlns:h="<http://www.w3.org/1999/xhtml>">

 <its:externalResourceRefRule selector="//h:img" externalResourceRefPointer="@src"/>

 </its:rules>

 </script>

 </head>

 <body>

 <p>Image: <img src="example.png" alt="Text for the image"></p>

 </body>

</html>

Extraction:

...

<trans-unit id='3'>

 <source>Image: <x id='1' itsxlf:externalResourceRef="example.png"/></source>

</trans-unit>

...

**Localization Quality Rating (==========TODO)**

Expresses an overall measurement of the localization quality of a document or an item in a document.
See <http://www.w3.org/TR/its20/#lqrating> for more details.

**Structural Elements**

Use the ITS attributes to annotate a <file>, <group>, <trans-unit> or <alt-trans> elements.

<trans-unit id="1" its:locQualityRatingScore="100"

 its:locQualityRatingScoreThreshold="95"

 its:locQualityRatingProfileRef="<http://example.org/qaModel/v13>">

 <source>text</source>

 <target>texte</target>

<trans-unit>

**Inline Elements**

Use the ITS attributes of Localization Quality Rating to annotate a segment (<mrk mtype="seg">) or a given span of the content (<mrk mtype="x-its">).

<trans-unit id="1">

 <source>Some text</source>

 <seg-source><mrk mtype="seg" mid="1">Some text</mrk></seg-source>

 <target><mrk mtype="seg" mid="1"

 its:locQualityRatingScore="0.56"

 its:locQualityRatingScoreThreshold="95"

 >Du texte</mrk></target>

</trans-unit>

<trans-unit id="1">

 <source>Some text and a term</source>

 <target>Du texte et un <mrk mtype="x-its" its:locQualityRatingVote="100"

 its:locQualityRatingVoteThreshold="95"

 its:locQualityRatingProfileRef="<http://example.org/qaModel/v13>">terme</mrk></source>

</trans-unit>

**Storage Size (==========TODO)**

Specifies the maximum storage size of a given content.
See <http://www.w3.org/TR/its20/#storagesize> for more details.

**Structural Elements**

Use the ITS attributes for Storage Size on the <source> and <target> elements:

<trans-unit id="1">

 <source its:storageSize="12"

 its:storageEncoding="UTF-16"

 its:lineBreakType="crlf">Text</source>

</trans-unit>

**Inline Elements**

Use the ITS attributes for Storage Size on the <mrk> element:

<trans-unit id="1">

 <source><mrk its:storageSize="8"

 its:storageEncoding="UTF-16" mtype="x-its">CONTINUE</mrk></source>

</trans-unit>

**References**

* Internationalization Tag Set (ITS) Version 2.0: <http://www.w3.org/TR/its20/>
* XLIFF Version 2.0: <http://docs.oasis-open.org/xliff/xliff-core/v2.0/xliff-core-v2.0.html>

**Navigation menu**

* [Log in](http://www.w3.org/International/its/wiki/index.php?title=Special:UserLogin&returnto=XLIFF+2.0+Mapping)
* [Page](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping)
* [Discussion](http://www.w3.org/International/its/wiki/index.php?title=Talk:XLIFF_2.0_Mapping&action=edit&redlink=1)
* [Read](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping)
* [View source](http://www.w3.org/International/its/wiki/index.php?title=XLIFF_2.0_Mapping&action=edit)
* [View history](http://www.w3.org/International/its/wiki/index.php?title=XLIFF_2.0_Mapping&action=history)

Top of Form



Bottom of Form

* [Main page](http://www.w3.org/International/its/wiki/Main_Page)
* [Recent changes](http://www.w3.org/International/its/wiki/Special%3ARecentChanges)
* [Random page](http://www.w3.org/International/its/wiki/Special%3ARandom)
* [Help](https://www.mediawiki.org/wiki/Special%3AMyLanguage/Help%3AContents)
* [Help](https://www.mediawiki.org/wiki/Special%3AMyLanguage/Help%3AContents)

[**Tools**](http://www.w3.org/International/its/wiki/XLIFF_2.0_Mapping)

* [What links here](http://www.w3.org/International/its/wiki/Special%3AWhatLinksHere/XLIFF_2.0_Mapping)
* [Related changes](http://www.w3.org/International/its/wiki/Special%3ARecentChangesLinked/XLIFF_2.0_Mapping)
* [Special pages](http://www.w3.org/International/its/wiki/Special%3ASpecialPages)
* [Printable version](http://www.w3.org/International/its/wiki/index.php?title=XLIFF_2.0_Mapping&printable=yes)
* [Permanent link](http://www.w3.org/International/its/wiki/index.php?title=XLIFF_2.0_Mapping&oldid=1051)
* [Page information](http://www.w3.org/International/its/wiki/index.php?title=XLIFF_2.0_Mapping&action=info)
* This page was last modified on 8 October 2014, at 12:02.
* This page has been accessed 7,466 times.
* [Privacy policy](http://www.w3.org/International/its/wiki/ITS%3APrivacy_policy)
* [About ITS](http://www.w3.org/International/its/wiki/ITS%3AAbout)
* [Disclaimers](http://www.w3.org/International/its/wiki/ITS%3AGeneral_disclaimer)
* [Mobile view](https://www.w3.org/International/its/wiki/index.php?title=XLIFF_2.0_Mapping&mobileaction=toggle_view_mobile)
* 