

# THE DOUBLE-EDGED SWORD OF EXTENSIBILITY IN XLIFF

YAMAGATA EUROPE | Thomas Vackier | September 22, 2010



WTF?

# EXTENSIBILITY

*Add non-XLIFF elements, as well as  
attributes and attribute values*

# Namespaces

Adding elements and attributes can be done using the namespace mechanism.

There are several extension points; `<source>` and `<target>` elements are **not** extensible.

```
<xliff version='1.2'
xmlns='urn:oasis:names:tc:xliff:document:1.2'
xmlns:foo='http://www.foo.com/BAR' >
. . . .
<trans-unit id='1' >
<source>Namespace</source>
<target>名前空間</target>
<foo:BlahBlah>Waffle</foo:BlahBlah>
</trans-unit>
. . . .
```

# Caveat

*Although XLIFF offers this extensibility mechanism, in order to avoid a nimety of information and increase interoperability between tools, it is strongly recommended to use XLIFF capabilities whenever possible, rather than to create non-standard user-defined elements or attributes.*

← Implementer beware!

**Caveat: implementer beware**

# Double-edged sword

## BENEFITS

*Flexibility*

*Enhance standard*

## RISKS

*Alternative approaches to standard methods*

*Non-documented extensions*

*Lagging official standard*

*XLIFF flavours*

*Watered-down standard*

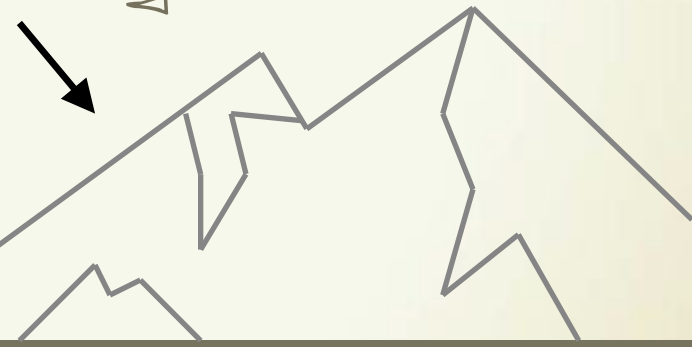
**Double-edged sword: benefits and risks**

# XLIFF flavours

Extensibility feature has been used to adapt & enhance XLIFF.

This has led to an abundance of custom XLIFF flavours with different namespaces.

*Tip of the iceberg!*

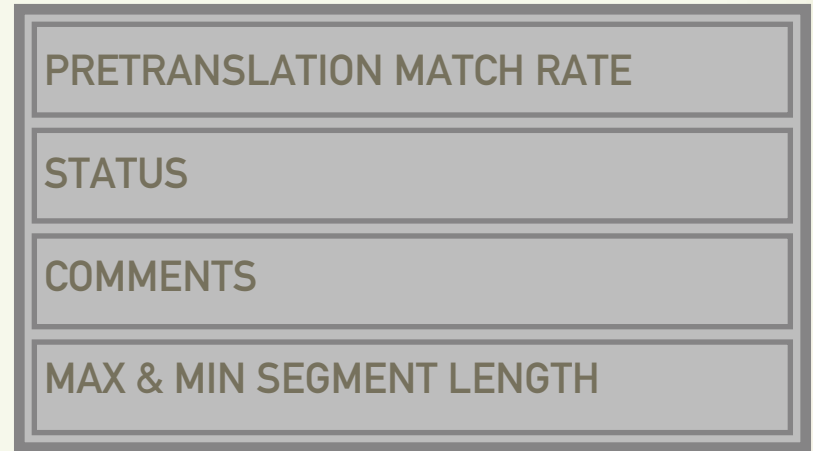


**XLIFF flavours:**



# XLIFF support

Supporting XLIFF fully means supporting extensions that are often used to carry valuable metadata.



*It's not only about the source and target*

**XLIFF support**

# QA Distiller

Detects measurable translation mistakes in bilingual files.

Supported formats are TTX, TMX and TBX for terminology.

XLIFF recently added.

Source and target have a different end of sentence punctuation

Source I have a book.

Target Tá leabhar agam

XLIFF support in QA Distiller

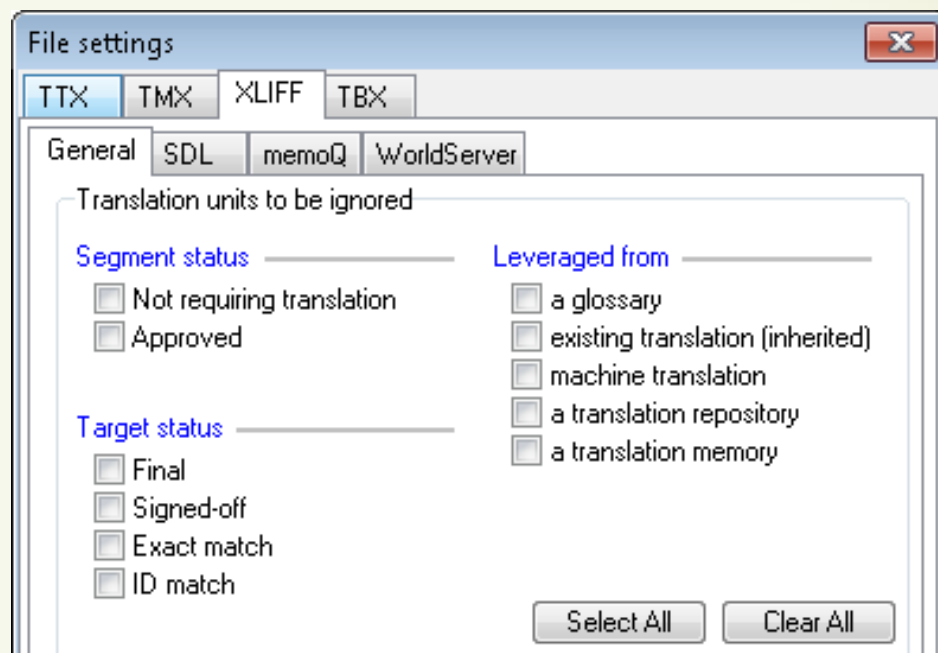




# QA Distiller uses metadata

Uses metadata to filter out or report certain translation units.

Similar information retrieved from different namespace elements or attributes.

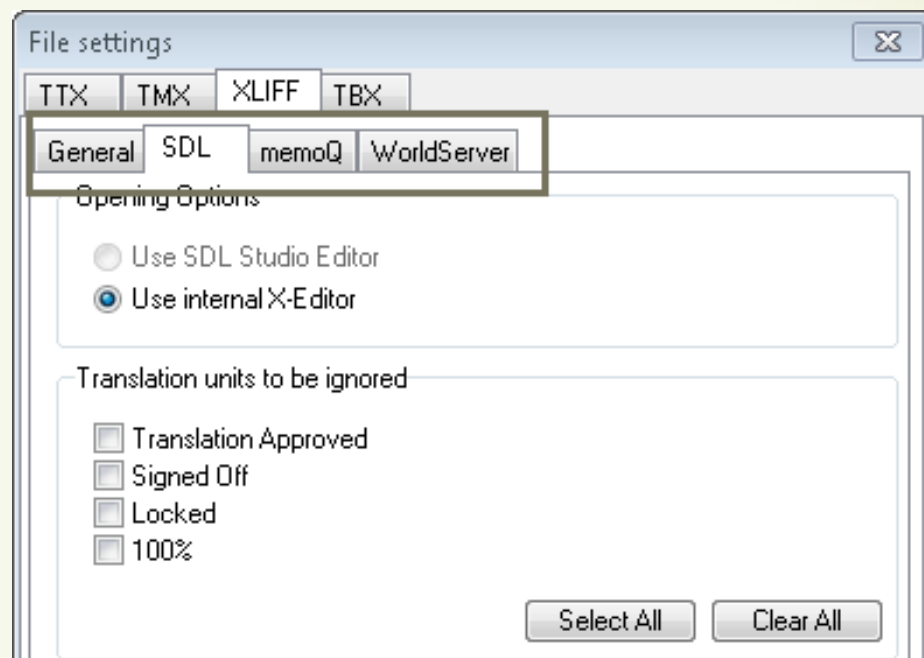


**XLIFF support in QA Distiller: metadata**



# QA Distiller XLIFF settings

Instead of a single place for all XLIFF settings, we needed to have different tabs to cater for the different flavours.



**XLIFF support in QA Distiller: XLIFF settings**



# NITTY GRITTY DETAILS

*Extensibility on the loose*

# Pretranslation match rate

Match rate qualifies match of  
<alt-trans>.

Only tool we found that stores match  
rate the standard way is Swordfish.

. . . .

```
<trans-unit id=' 1' >
<source xml:lang=' en' >Whi te cat</source>
<target xml:lang=' nl ' >Wi tte kat</target>
<al t-trans match-qual i ty=' 85' >
<source xml:lang=' en' >Whi te dog</source>
<target xml:lang=' nl ' >Wi tte hond</target>
</al t-trans>
</trans-unit>
```

. . . .

Pretranslation match rate: the standard way



# Pretranslation match rate

**MEMOQ**

*//trans-unit/@mq:percent*

**SDL**

*//trans-unit/sdl:seg-defs/sdl:seg/@percent*

**WORLDSEVER**

*//trans-unit/iws:segment-metadata/@tm\_score*

**Pretranslation match rate: every other way**

# State

The standard provides two attributes for the `<target>` element: *state* and *state-qualifier*.

There are pre-defined values but they also take user-defined values.

Well anticipated!



```
. . . .  
<trans-unit id=' 1' >  
<source>Dog</source>  
<target state=' signed-off' >Hond</target>  
</trans-unit>  
. . . .
```

State: the standard way



# State

They've gone and done it again!



**MEMOQ**

*//trans-unit/@mq:status*

**SDL**

*//trans-unit/sdl:seg-defs/sdl:seg/@conf*

**WORLDSEVER**

*//trans-unit/iws:segment-  
metadata/iws:status/@translation\_status*

**State: every other way**

# Commenting

The `<note>` element can be inserted at different positions and for `<trans-unit>` elements you can specify if the comment relates to the `<source>` or `<target>` element.

```
. . . .  
<trans-unit id=' 1' >  
  <source xml:lang=' en' >Namespace</source>  
  <target xml:lang=' nl ' >Naamruimte</target>  
  <note annotates=' target' >Waffle waffle</note>  
</trans-unit>  
  
. . . .
```

**Commenting: standard way**



# Enhanced commenting

SDL XLIFF uses the `<mrk>` element with the *mtype* attribute set to *x-sdl-comment*.

This allows for comments to be associated with content inside `<source>` or `<target>` elements.

```
. . . .  
<target>  
<mrk mtype="seg" mi d="22">  
<mrk mtype="x-sdl -comment"  
sdl : ci d="00A">Update</mrk> und ei ne Warnung  
</mrk>  
</target>  
. . . .
```



# THE FUTURE

**IMPLEMENTATION**  
*trumps*  
**SPECIFICATION**

*Get in Touch!*

**@YamagataEurope**  
**@QADistiller**