

MultilingualWeb-LT Working Group

5 min 30 sec overview

@ TPAC 2012

Felix Sasaki

Co-chair and staff contact

Internationalization Tag Set (ITS) 2.0

- Defining metadata for multilingual processing of Web or other content, e.g.
 - Machine translation
 - Localization workflows
- Where is the metadata needed – *for example*:
 - In Web content, e.g. HTML5
 - In XML-based formats like DocBook etc.
 - In Localization related formats like XLIFF (XML Localization Interchange File Format)

Metadata example: “Translate” in HTML5 and XLIFF

```
<!DOCTYPE html>
<html> ...
<p>The <span translate=no>World Wide Web Consortium</
span> is making the World Web Web worldwide!</p>...</html>
```

```
<xliff ...> ...
  <trans-unit id="1">
    <source xml:lang="en">The <mrk mtype="protected">World
Wide Web Consortium</mrk> ...!</source>
    <target> ...
  </xliff>
```

Metadata example: “Terminology” in HTML5 and XLIFF

```
<!DOCTYPE html>  
<html> ...  
<p>We need a new <span its-term=yes>motherboard</span>  
...</html>
```

```
<xliff ...> ...  
  <trans-unit id="1">  
    <source xml:lang="en">We need a new  
      <mrk mtype="term">motherboard</mrk></source>  
    <target> ...  
  </xliff>
```

More about ITS 2.0 – e.g.:

- More metadata items (“data categories”)
 - See sec. 6 of <http://www.w3.org/TR/its20/>
- Mechanisms to attach metadata to several nodes
 - Using mostly XPath, see <http://www.w3.org/TR/its20/#EX-translate-html5-global-1>
- Applications
 - See demo in 2 min

Plan and goal today

- Move 23 October draft to last call by end of November
 - <http://www.w3.org/TR/its20/>
- Have HTML5+ITS schema in HTML validator
 - See current offline validation at <https://github.com/kosek/html5-its-tools>
- Get feedback from HTML WG on “ITS 2.0 in HTML5 section” by end of November
 - <http://www.w3.org/TR/its20/#html5-markup>
- Today: Show a demo & get feedback
 - On above time schedule, in general etc.