

FEISGILTT 2016 XLIFF Public Info sessions

Draft meeting minutes

Tue June 7 2016

Participants:

[participation in particular sessions was not tracked, this list is based on FEISGILTT 2016 participation]

XLIFF TC members

Chair

N/A

Secretary

David Filip (ADAPT Centre)

Members

Ryan King (Microsoft), Kevin O'Donnell (Microsoft), Soroush Saadatfar (ADAPT Centre @ LRC @ University of Limerick), Felix Sasaki (Individual), Phil Ritchie (Vistatec),

Public

Victor Alves (Individual), Jan Bareš (Moravia), George Bina (Syncro Soft), Konrad Chmielewski (XTRF), David Clarke (Welocalize), Mariza Flores (Thousand Words Translations), Andreas Galambos (Transmission Übersetzungne), Andrew Gibbons (Welocalize), Ján Husarčík (Moravia), Łukasz Kaleta (XTRF), Gary Lefman (Cisco), David Lewis (ADAPT Centre), Ian Morris (Individual), Dimitris Orfanoudakis (TAUS), Achraf Oumghar (Lionbridge), Felix Sasaki (DFKI), Uta Seewald-Heeg (Anhalt University of Applied Sciences), Richard Sikes (Content Rules), Vinod Sudharshan (TAUS), Chase Tingley (Spartan Software), Angelika Zerfaß (zaac).

XLIFF TC Public Info Sessions started at 10.15am WEST on June 7th 2016.

XLIFF 2.1 Work in Progress: ITS Module

Chair: dF

Panelists: XLIFF TC membership

This was introduced by a presentation by Soroush and dF on what it means for ITS to become a module rather than an extension as of XLIFF 2.1 <http://lanyrd.com/2016/feisgiltt/sfbtzt/>

During the info session, the TC members informed the public on what is the current state of the ITS module development. The audience and TC members expressed interest mainly in the Text Analytics and Provenance metadata categories

XLIFF 2.1 Work in Progress: Module Fixes

Chair: dF

Panelists: XLIFF TC membership

dF inquired with Phil about the state of the Change Tracking module fix that is being led by Phil. Phil said there was unfortunately no progress on that so far. dF said that the module fix should be done by the end of August, should it be included in the 2.1 release as the ITS and Advanced Validation features are near completion and the 1st public review shouldn't start later than in August 2016.

Other module fixes should include explicitly allowing the metadata module on most of the other modules. This is now explicitly allowed only on the Translation candidates module.

XLIFF 2.1 Work in Progress: Advanced Validation

Chair: dF

Reporters: Soroush Saadatfar (ADAPT Centre @ LRC @ University of Limerick) & Felix Sasaki (DFKI)

Soroush and Felix demonstrated some of the Advanced XLIFF Constraints that are not expressible using XML Schema.

dF explained that the Advanced Constraints had be validated by custom code in the available XLIFF 2.0 validators such as OKAPI Lynx <http://okapi-lynx.appspot.com/validation>, Xmarker <http://xmarker.com/node/13>, or the Microsoft XLIFF Object Model <https://github.com/Microsoft/XLIFF2-Object-Model>

XLIFF TC Public Info Sessions broke for lunch at 12:30 WEST.

XLIFF TC Public Info Sessions resumed at 16:00 WEST [after lunch and XLIFF OMOS Info Sessions].

XLIFF 2.2 Requirements Gathering

Chair: dF

Reporter: Chase Tingley

Panel: All

Plurals?

Chase reported on a plurals support discussion that took place on the OKAPI developer forum.

dF reminded Chase and others that there was a discussion about plurals support in the early stages of XLIFF 2.0 development. David Walters from IBM said then that we don't need to support the old limited notion of plurals and that now conditional plural structures are part of CLDR data.

dF also reported on a ULI discussion that basically agreed that the proper I18n approach is to have different strings based on the CLDR locale specifics.

Chase and dF discussed how this could be handled during the XLIFF roundtrip either on group or subunit level. The advantage of the subunit handling would be that the modifiers would be able to create the conditional strings w/o violating XLIFF structural Constraints. The subunit conditional segments should be protected from re-segmenting so that agents not understanding the plural handling module couldn't break the feature.

In either case new module metadata would need to be developed to describe the conditional multiplicity groups and placed either on the unit or the group extension point.

dF suggested that this should be proposed as a 2.2 feature by roughly end of October 2016 and that Chase better joined the XLIFF TC so suggest that. Otherwise, this would need to be suggested over the commenting facility and Chase couldn't drive the feature.

Segmentation information?

Including segmentation rule information in XLIFF 2.2 was also discussed. That was suggested by ULI representatives in Berlin 2015. However, no one proposed the feature to XLIFF TC formally. dF suggested that this could be done via inclusion of SRX as module data. Chase argued that not all rule were SRX expressible. A short discussion revealed that what Chase meant that statistically driven methods of segmentation cannot be described with rules. dF suggested that rule based segmentation should be always encoded as SRX, whereas there should be an option record just provenance of the tool if the method wasn't rules driven.

dF also informed that the ULI stakeholders would be in favor of using the BCP47 t-extension to roundtrip this data via XLIFF. Nevertheless, dF recalled a discussion with Mark Davis (video presence) during ITS

2.0 Requirements gathering in Dublin 2012 http://videolectures.net/w3cworkshop2012_davis_data/, where Davis explicitly said that the t and u extensions were not suitable for structured XML content that should store this metadata using its own markup. dF said that one of the main issues of the extension based segmentation information that it was not stateful and it wouldn't be possible to tell if the method was applied, had to be applied, if it was preserved etc.

This could also be formally submitted to XLIFF TC as a feature for XLIFF 2.2 consideration.

Register Information?

Ján Husračík from Moravia suggested that language register variants (polite, informal, archaic, intimate etc.) should be also treated in XLIFF at some point. dF explained that this would need to be first codified for various locales in the CLDR. XLIFF as a transfer format would overstep its scope if it tried to codify locale specifics in this area in a material way. dF reported on a similar discussion that took place during the W3C ITS 2.0 Requirements Gathering. The ITS 2.0 stakeholders agreed back in 2012 that this would need to be addressed at the CLDR level first.

XLIFF TC Public Info Sessions were adjourned on 7th June 2016 at 5:15pm WEST