

ILO Extension User Guide

Sebastian Sklarß, Matthias Kandora

© 2013]init[AG für digitale Kommunikation Berlin,

The extension ILO including this documentation was published as LGPL v3 in August 2013 as part of the EU funded MultilingualWeb-LT initiative, Grant Number 287815

<http://www.multilingualweb.eu/about-the-project>



Table of content

Overview.....	3
Installation.....	3
Basic Usage.....	3
Insert a Data Category.....	4
Modify An Existing Data Category.....	5
Removing An Existing Data Category.....	6
Exporting to XLIFF.....	7
Importing from XLIFF.....	7

Overview

The ILO extension offers the possibility to enrich documents with the second version of the W3C Internationalization Tag Set (ITS 2.0)¹. This W3C standard offers the possibility to add so called ITS data categories on content. This metadata markup is foreseen to ease up processes between content providers, translation agencies and machine translation systems. This extension implements four basic tags: The *translate*² tag, to indicate a paragraph (or a spanning of text), that should be translated or not, the *locale filter*³ tag to add or restrict information only to a specific region, a *localization note*⁴ to pass additional information to the localization service provider and a *terminology*⁵ annotation to markup certain words as terms.

The extension offers support for the XLIFF (XML Localization Interchange File Format), that can be used by translation agencies and machine translation systems.

Installation

The installation of the ILO extension is easy and straightforward. Open LibreOffice, go to the extension manager located under “Extras > extension manager“ use “add extension“ and move to the folder, where you downloaded the ILO extension, select it and it will be installed automatically after you opted in for the LGPL license. **Libre Office has to be restarted to take into account the new extension.** After the installation a small toolbar will appear on LibreOffice Writer's Screen giving you access to the new ITS functionality:



Basic Usage

This chapter explains some of the basic operations that can be done using the extension. The following instructions will explain how to mark up content with a specific data category, how to change previously configured data categories and how to delete them.

Insert a Data Category

Consider you wrote some text, that contains certain passages, that should not be translated by a translation agency. Reasons for this might be that it contains some citations that have to remain in original language or phrases that are even untranslatable. To markup this passage as a non-translatable to the translation agency you can use the translate button from the ILO toolbar and select “no”. Follow these steps:

1. Markup the passage you want to mark as non-translatable:

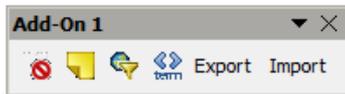
¹ <http://www.w3.org/TR/its20/>

² <http://www.w3.org/TR/its20/#trans-datacat>

³ <http://www.w3.org/TR/its20/#LocaleFilter>

⁴ <http://www.w3.org/TR/its20/#locNote-datacat>

⁵ <http://www.w3.org/TR/its20/#terminology>



Lorem ipsum dolor sit amet, consectetur adipiscing elit. Aenean commodo ligula eget dolor. Aenean massa. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. **Donec quam felis, ultricies nec, pellentesque eu, pretium quis, sem.** Nulla consequat massa quis enim. Donec pede justo, fringilla vel, aliquet nec, vulputate eget, arcu. In enim justo, rhoncus ut, imperdiet a, venenatis vitae, justo. Nullam dictum felis eu pede mollis pretium. Integer tincidunt. Cras dapibus. Vivamus elementum semper nisi. Aenean vulputate eleifend tellus. Aenean leo ligula, porttitor eu, consequat vitae, eleifend ac, enim. Aliquam lorem ante, dapibus in, viverra quis, feugiat a, tellus. Phasellus viverra nulla ut metus varius laoreet. Quisque rutrum. Aenean imperdiet. Etiam ultricies nisi vel augue. Curabitur ullamcorper ultricies nisi. Nam eget dui. Etiam rhoncus.

2. Use the first toolbar icon to display the translate data category dialog:



3. Select “yes“, to mark the selection as translatable or “no“ to mark it as non-translatable. After you have successfully decided on one of the two options, the content gets a specific style to optically highlight the existence of the translate tag:



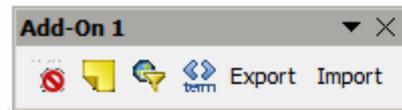
Lorem ipsum dolor sit amet, consectetur adipiscing elit. Aenean commodo ligula eget dolor. Aenean massa. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. **Donec quam felis, ultricies nec, pellentesque eu, pretium quis, sem.** Nulla consequat massa quis enim. Donec pede justo, fringilla vel, aliquet nec, vulputate eget, arcu. In enim justo, rhoncus ut, imperdiet a, venenatis vitae, justo. Nullam dictum felis eu pede mollis pretium. Integer tincidunt. Cras dapibus. Vivamus elementum semper nisi. Aenean vulputate eleifend tellus. Aenean leo ligula,

Notice: Even though it is possible to nest multiple data categories of the same kind, a technical limitation is set to implement a useful styling for nested data categories. A section can contain more than one data category but the existing hierarchy may not be visible (you will still see it as a text mark in the libre office navigator).

Modify An Existing Data Category

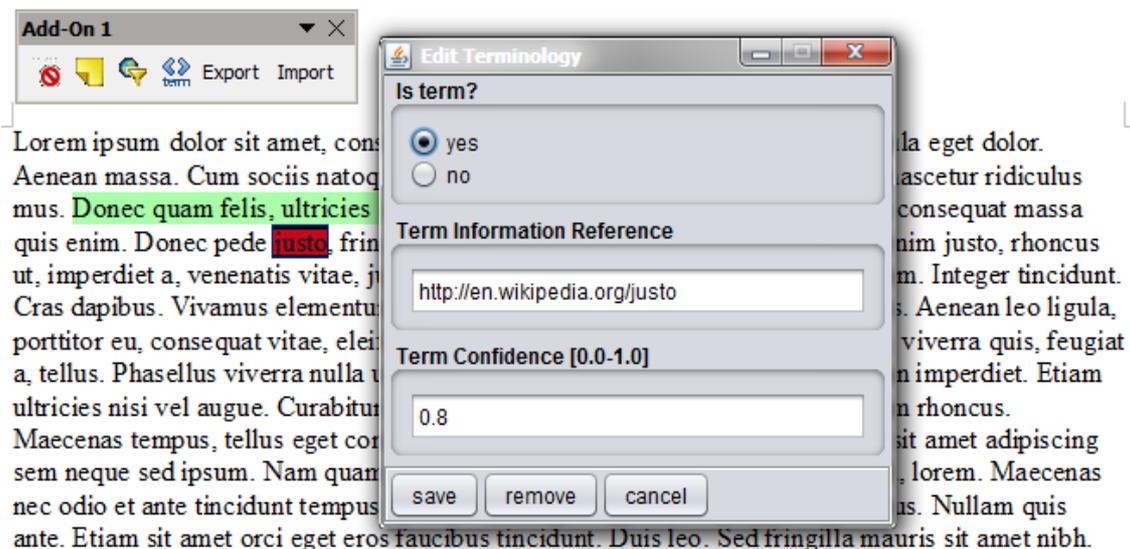
After you have marked up some content with different data categories, you may want to change the values you have configured earlier. To do so, select a marked section and select the appropriate data category to display the configured values:

1. Make a selection (The screenshot has the word 'justo' selected):



Lorem ipsum dolor sit amet, consectetur adipiscing elit. Aenean o
 Aenean massa. Cum sociis natoque penatibus et magnis dis parturi
 mus. Donec quam felis, ultricies nec, pellentesque eu, pretium quis
 quis enim. Donec pede **justo**, fringilla vel, aliquet nec, vulputate eg
 ut, imperdiet a, venenatis vitae, justo. Nullam dictum felis eu pede

2. Now select the appropriate data category (in this example we chose 'terminology') to display the dialog:

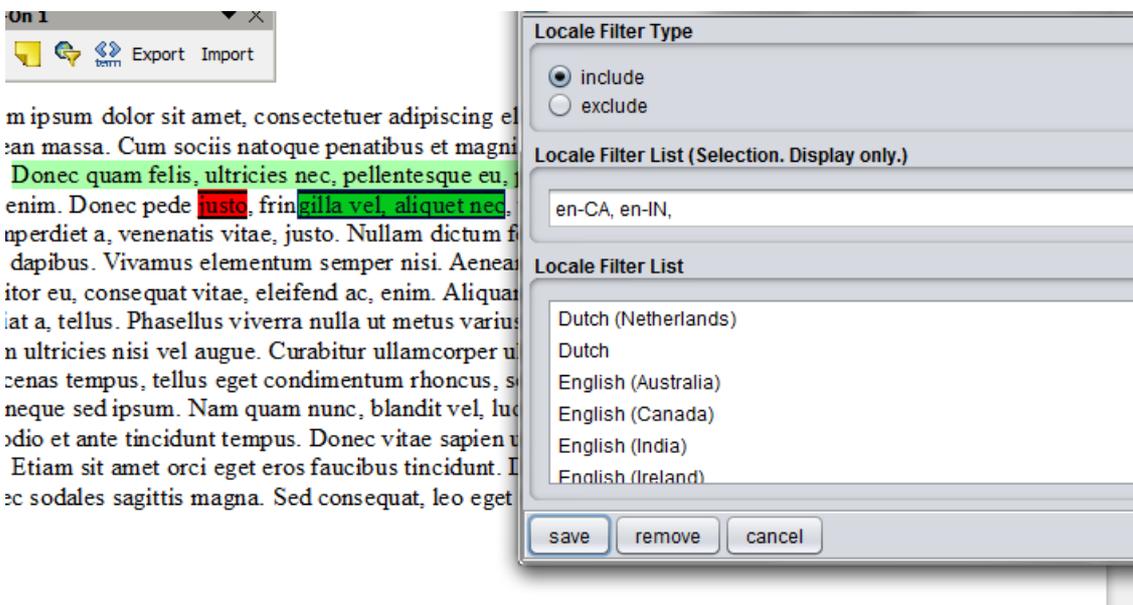


As you can see in the screenshot, 'justo' is a term with a confidence of 80% and contains an information reference to define this term.

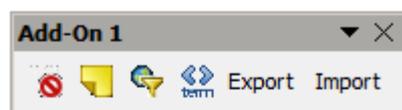
Removing An Existing Data Category

Sometimes some passages you marked with a specific data category are simply wrong and should not contain any meta information. To remove a specific data category make an exact selection and select the appropriate data category from the toolbar:

1. select the content marked with a specific data category and open the dialog from the toolbar (in this example 'gilla vel, aliquet nec' is selected)



2. select 'remove' from the dialog to remove the data category from the selection.



Lorem ipsum dolor sit amet, consectetur adipiscing elit. Aenean comm
 Aenean massa. Cum sociis natoque penatibus et magnis dis parturient m
 mus. Donec quam felis, ultricies nec, pellentesque eu, pretium quis, sem
 quis enim. Donec pede justo, fringilla vel, aliquet nec, vulputate eget, ar
 ut, imperdiet a, venenatis vitae, justo. Nullam dictum felis eu pede molli
 Cras dapibus. Vivamus elementum semper nisi. Aenean vulputate eleifend

3. the data category has successfully been removed!

Exporting to XLIFF

The ILO extension supports the extraction of textual content with ITS markup. Once you have finished editing your document, configured the values of every data category you can export your document including the ITS information to the XLIFF format by pressing the Export Button located on the toolbar.

Importing from XLIFF

The ILO Extension also supports as kind of a roundtrip the import of XLIFF documents, without the loss of any textual information. Use the import button on the right corner of the toolbar to import an existing XLIFF document into LibreOffice. A new document window will appear.