

a folha

Boletim da língua portuguesa nas instituições europeias

<http://ec.europa.eu/translation/portuguese/magazine>

English version of an article published in No. 54 — Summer 2017

Free open source software – CAT Tools: OmegaT in DGT, its Wizard, Tagwipe and TeamBase (Part 1)

*Maria José Bellino Machado; Hilário Leal Fontes - Portuguese Language Department
Thomas Cordonnier; Elio Fedele; Fons De Vuyst – Informatics Unit
Directorate-General for Translation — European Commission*

[click [here](#) for Portuguese version]

1. Introduction

In a previous article in ‘a folha’ published in 2014⁽¹⁾, the use of free open source software in DGT – namely the Computer-Assisted Translation tool OmegaT⁽²⁾ – was presented in general terms.

In this article, which is divided in 2 parts, the following Computer-Assisted Translation (CAT) applications which have been recently made available by the Directorate-General for Translation (DGT) of the European Commission as free open source software (FOSS) are presented:

- **OmegaT in DGT** (which we call DGT-OmegaT) with adaptations/improvements/new features of DGT’s implementation of OmegaT (standard version 3.6⁽³⁾) developed in-house to satisfy DGT’s needs (Developer: Thomas Cordonnier)
- **DGT-OmegaT Wizard** (Developer: Elio Fedele)
- **Tagwipe** (Developers: Elio Fedele and João Rosas)
- **Teambase** (Developer: Thomas Cordonnier)

The project is managed by Fons De Vuyst, Head of the Operational Support Sector in DGT’s Informatics Unit.

In the first part of this article, we present a brief description of: a) DGT’s workflow in the context of which these applications have been adapted/developed; b) Tagwipe; and c) 10 features adapted/added to OmegaT.

In the second part of this article, which will be published at the end of 2017, we will present the DGT-OmegaT Wizard and TeamBase, two applications which still require some work to be “usable” outside DGT.

Concerning the publication of DGT-OmegaT as open source software, we would like to stress that our aim is not to build a new community that "competes" with the team developing and maintaining the original OmegaT. DGT publishes the code just to allow anyone to look at the features we have built to

satisfy our own needs and leave it up to the OmegaT community to decide whether some of them are of general interest and can be reused and integrated in it.

These applications can be downloaded and easily tried/used by anyone in any Java8-compliant platform (Windows, Linux, MacOS), as both the source code and the executable version are published, but bearing in mind that they are published as-is, without express or implied warranties, and come without support.

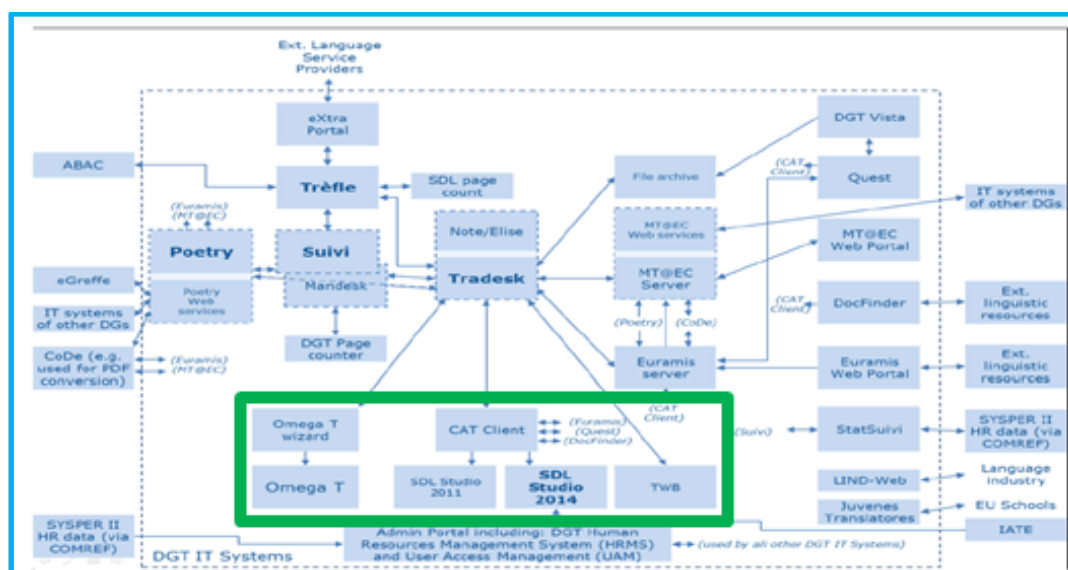
2. Open source software in the European Commission

The European Commission has been promoting the use — and sharing — of open-source software namely within its Framework Programmes on Research and Technological Development — 7th FP and presently Horizon 2020⁽⁴⁾. Furthermore, the Open Source Software Strategy of the European Commission⁽⁵⁾ stresses that “*the Commission services will increasingly participate in open source software communities to build on the open source building blocks which are used in the Commission's software*”.

So it is in this context that open source applications like OmegaT are being used in DGT, alongside commercial applications like SDL Trados Studio, the mainstream CAT tool in DGT.

DGT has been using OmegaT for prototyping since 2012 and for that purpose the 2.6.0._3 version of OmegaT was first customized and extended with some useful improvements.

Some translators liked OmegaT so much that it has been included in **DGT's IT Landscape** and for that reason DGT-OmegaT Wizard was developed to integrate OmegaT in DGT's particular workflow.



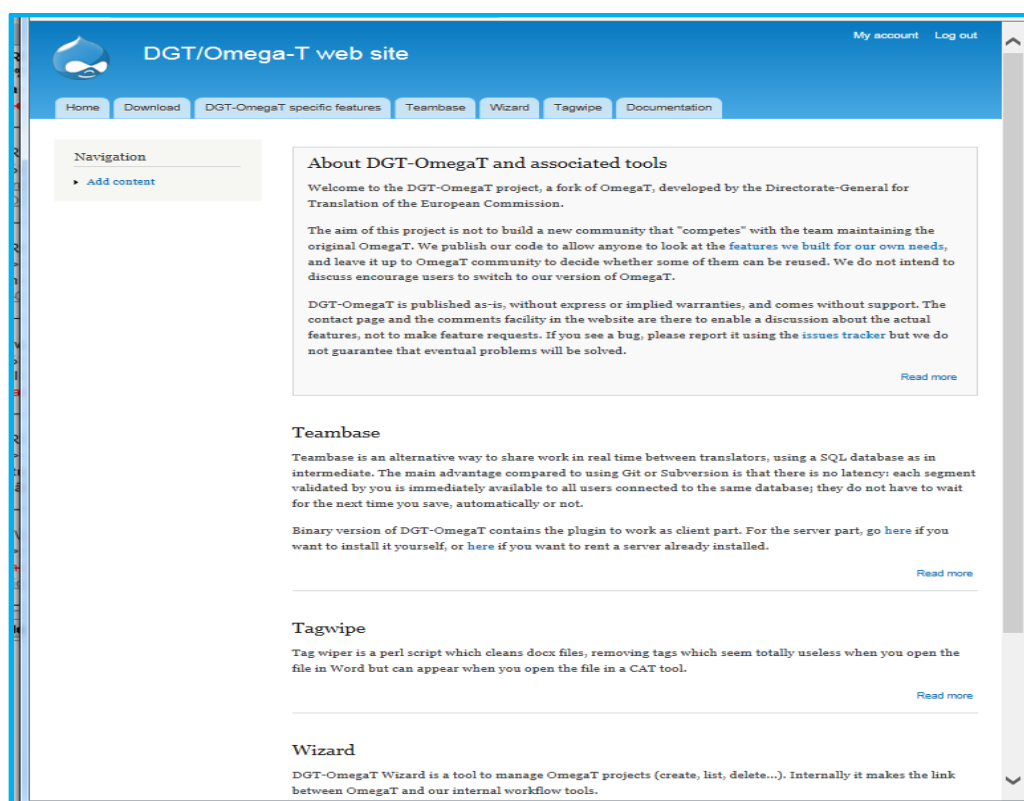
DGT-IT LANDSCAPE

Directorate-General for Translation, *Information Technology Master Plan 2016–2020*

Furthermore, 2 other applications were developed, which can be used both by OmegaT and SDL Studio users:

- TagWipe, an application which removes redundant tags inside Office DOCX files.
- TeamBase, which allows the sharing of project memories among translators who can be using either of the CAT tools available in DGT

Recently, DGT made these applications available as free open source software and they can be downloaded here: <http://185.13.37.79/>.



3. DGT Workflow

DGT's CAT Environment is explained in detail in the DGT publication *Translation tools and workflow*⁽⁶⁾.

Of importance for the purposes of this article is that, in DGT:

- **Translation memories** relevant for the particular document(s) to be translated are automatically extracted from Euramis, the in-house repository which stores all the alignments of translations done in DGT — as well as in some other EU institutions — in the last 2 decades and most of the EU legislation currently in force in all official languages. Retrievals and reference documents are extracted in TMX format and used as external memories in either of the CAT tools. In DGT-OmegaT, those memories are automatically copied to the `\tm` subfolder of each project by the DGT-OmegaT Wizard.
- **Machine translation** is automatically generated for all editable documents to be translated. This task is performed by MT@EC⁽⁷⁾, the in-house machine translation system based on the free open source system Moses⁽⁸⁾, which is trained with EU corpora and tailored to DGT needs. The TMX files generated by MT@EC are used as external memories in either of the CAT tools. In DGT-OmegaT, those files are automatically copied to the `\mt` subfolder of each project by the DGT-OmegaT Wizard.
- **Terminology** is stored in IATE (Inter-Active Terminology for Europe) – the EU inter-institutional terminology database. In DGT-OmegaT projects, the relevant terms for the particular document(s) to be translated are extracted in TXT format from an export (source and target terms only) of the whole IATE and this “filtered” glossary is automatically copied to the `\glossary` subfolder of each project by the DGT-OmegaT Wizard as a read-only glossary.

4. Tagwipe

More often than not Office DOCX documents (the huge majority of DGT documents) have (many) useless tags (the “tag soup”). This obviously makes it impossible to use OmegaT without **Remove tags** selected in a number of (unfortunately not so rare) particularly “bad” documents. However, that is not the best solution for obvious reasons.

<p>Many point to possible further improvements in the use of the ESIFs and call for NGOs and local authorities to be given direct access to funds, for better enforcement of <i>ex ante</i> conditionalities, sanctions for failure to uphold the partnership principle, better monitoring (through an increased role for the Commission and Roma themselves) and action to prevent the ineffective use of funds (e.g. training programmes not leading to employment) or their misuse (e.g. ESIF interventions financing segregated settings), including through a transparent complaint mechanism.</p>	<p><176/>Many <177/>point to<178/> <179/>possible further improvements <180/>in<181/> <182/>the<183/> use of<184/> <185/>the <186/>E<187/>SIF<188/>s<189/> and call for NGOs and local authorities<190/> <191/>to be given<192/> direct access to funds<193/>, <194/>for <195/>better enforcement of <196/>ex ante <197/>conditionalities, sanctions for <198/>failure to uphold <199/>the partnership principle, better monitoring (through<100/> an<101/> increased role <102/>for <103/>the Commission and Roma themselves) <104/>and <105/>action to <106/>prevent<107/> the ineffective use <108/>of funds <109/>(e.g.<110/><111/>training programmes not leading to employment) <112/>or<113/> <114/>their <115/>misuse <116/>(e.g.<117/><118/>ESIF interventions financing segregated settings)<119/>, including through<120/> a transparent complaint mechanism<121/>.</p>
---	--

Therefore, DGT developed in-house the Tagwipe application which removes all/most redundant tags from DOCX documents. It also improves segmentation by segment, something which is very important for consistency with segmentation in MT@EC and Euramis, thereby improving fuzzy matching with external memories.

In DGT, Tagwipe is being used since 2012, in a Windows 7 environment, for DGT-OmegaT projects and therefore it is quite stable. In the DGT-OmegaT website it is also made available for the Linux distribution Ubuntu.

The cleaning level in Tagwipe can be chosen by the user from level 0 to 8. By default, in the installation folder, the cleaning level defined is the second lowest and most conservative level ("**level.1**").

That is the level used in DGT as translated documents have to be generated with all the original formatting.

However, when that is not the case, the cleaning level can be increased, thereby eliminating non-essential formatting as, for instance, colour highlights, non-breaking spaces... or even (almost) all formatting when level 8 is selected.

Tagwipe is automatically used in the creation of translation projects for both CAT tools. In SDL Studio projects, tagwiping happens before the conversion of source files to SDLXLIFF format.

In DGT-OmegaT projects, Tagwipe has been integrated in the DGT-OmegaT Wizard so all DOCX source documents are tagwiped when the projects are created or updated.

In the example below, about 90% of the tags were eliminated by Tagwipe (from 121 tags to 14 tags) and segmentation was improved. The remaining tags are meaningful tags.

Display in OmegaT Editor		
	Without Tagwipe	With Tagwipe
<p>Other challenges and priorities largely overlap with the achievements. Stakeholders refer to declining levels of political commitment, with emerging priorities, such as the refugee crisis, shifting Member States' attention away from Roma inclusion. As regards reporting by Member States, they call for more transparency, further development and closer involvement of civil society in monitoring, which should also facilitate peer learning. They call for more attention to fighting discrimination, e.g. by launching more infringement proceedings and imposing sanctions on non-compliant Member States. Regret is expressed at the fact that the Commission's Roma-targeted and mainstream policy guidance under the EU framework and Europe 2020 is not enforceable. There are calls to target Roma more explicitly under European and national programmes, such as the youth guarantee and Erasmus+. Many point to possible further improvements in the use of the ESIFs and call for NGOs and local authorities to be given direct access to funds, for better enforcement of <i>ex ante</i> conditionalities, sanctions for failure to uphold the partnership principle, better monitoring (through an increased role for the Commission and Roma themselves) and action to prevent the ineffective use of funds (e.g. training programmes not leading to employment) or their misuse (e.g. ESIF interventions financing segregated settings), including through a transparent complaint mechanism.</p>	<p>Other challenges and priorities largely overlap with the achievements. Stakeholders refer to declining levels of political commitment, with emerging priorities, such as the refugee crisis, shifting Member States' attention away from Roma inclusion. As regards reporting by Member States, they call for more transparency, further development and closer involvement of civil society in monitoring, which should also facilitate peer learning. They call for more attention to fighting discrimination, e.g. by launching more infringement proceedings and imposing sanctions on non-compliant Member States. Regret is expressed at the fact that the Commission's Roma-targeted and mainstream policy guidance under the EU framework and Europe 2020 is not enforceable. There are calls to target Roma more explicitly under European and national programmes, such as the youth guarantee and Erasmus+. Many point to possible further improvements in the use of the ESIFs and call for NGOs and local authorities to be given direct access to funds, for better enforcement of ex ante conditionalities, sanctions for failure to uphold the partnership principle, better monitoring (through an increased role for the Commission and Roma themselves) and action to prevent the ineffective use of funds (e.g. training programmes not leading to employment) or their misuse (e.g. ESIF interventions financing segregated settings), including through a transparent complaint mechanism.</p>	<p>Other challenges and priorities largely overlap with the achievements. Stakeholders refer to declining levels of political commitment, with emerging priorities, such as the refugee crisis, shifting Member States' attention away from Roma inclusion. As regards reporting by Member States, they call for more transparency, further development and closer involvement of civil society in monitoring, which should also facilitate peer learning. They call for more attention to fighting discrimination, e.g. by launching more infringement proceedings and imposing sanctions on non-compliant Member States. Regret is expressed at the fact that the Commission's Roma-targeted and mainstream policy guidance under the EU framework and Europe 2020 is not enforceable. There are calls to target Roma more explicitly under European and national programmes, such as the youth guarantee and Erasmus+. Many point to possible further improvements in the use of the ESIFs and call for NGOs and local authorities to be given direct access to funds, for better enforcement of ex ante conditionalities, sanctions for failure to uphold the partnership principle, better monitoring (through an increased role for the Commission and Roma themselves) and action to prevent the ineffective use of funds (e.g. training programmes not leading to employment) or their misuse (e.g. ESIF interventions financing segregated settings), including through a transparent complaint mechanism.</p>

5. OmegaT in DGT

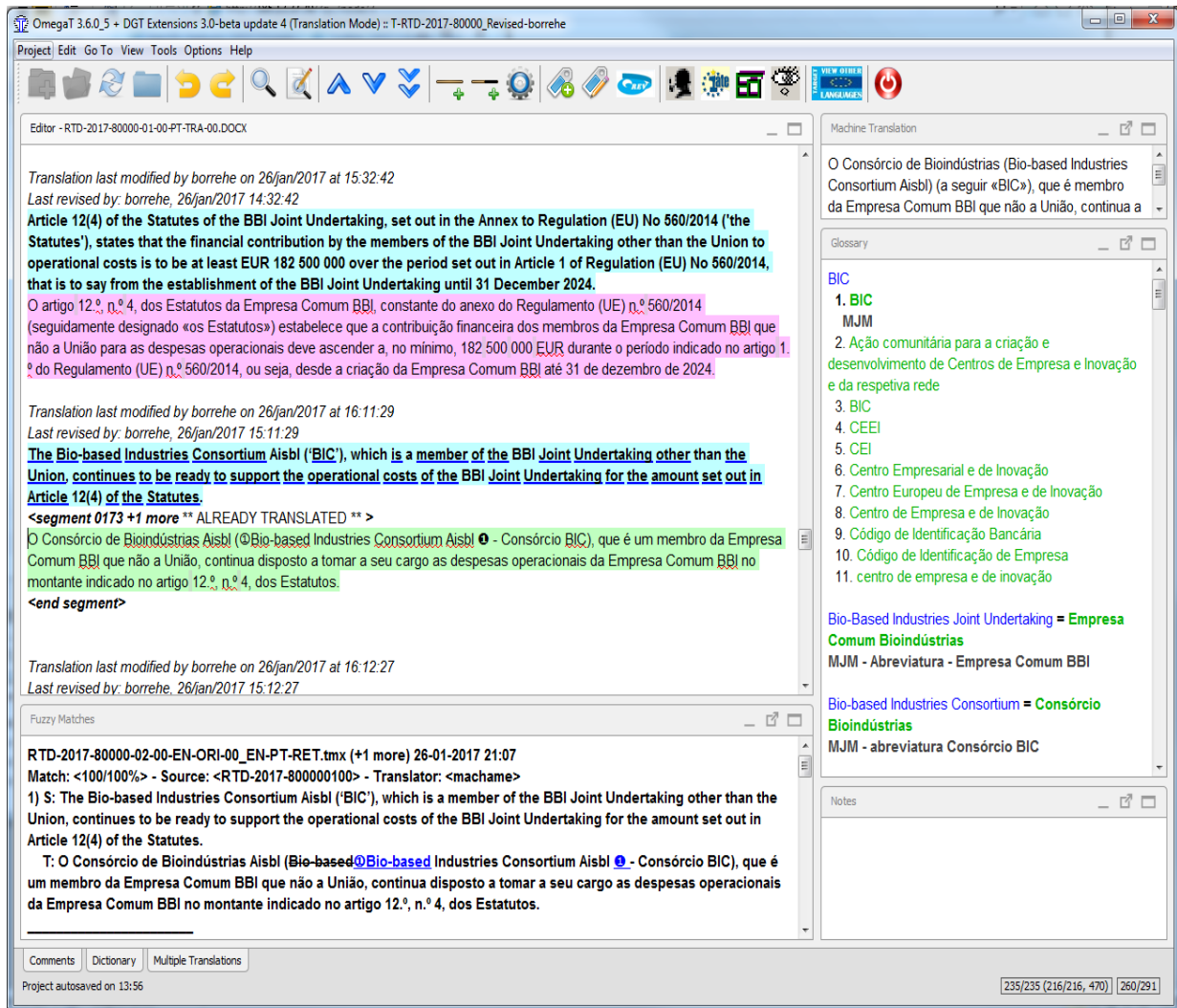
OmegaT is a free open-source CAT Tool that was developed, by private initiative, originally by Keith Godfrey in 2000 and that has been vastly improved since then with many contributions. Didier Briel is its present project manager. OmegaT is now the leading open source CAT tool.

OmegaT in DGT (DGT-OmegaT) is a fork of OmegaT 3.1.2 with a few backports from later versions. A new version based on OmegaT 3.6.0 and following the "**OmegaT standard**" branch is also available but as a beta development. Both can be downloaded here:

- <http://185.13.37.79/> to download the binary version (easy to install) and/or the source code and respective documentation
- <http://185.13.37.79:8003/> for developers and bug reporting

Despite the differences between DGT-OmegaT and OmegaT, as far as we can see, projects created in DGT-OmegaT can be used in OmegaT and vice-versa – as DGT-OmegaT maintains interoperability – and both just ignore folders/files/information they don't "recognise", although some features will not be available, of course.

In the DGT-OmegaT website detailed technical information is presented about each new or adapted feature. In this article we will just summarise/highlight 10 main features, from the point of view of a translator, which may be of interest for translators outside DGT.



5.1. DGT Toolbar

A **plugin** was developed in-house for DGT-OmegaT that produces a toolbar with icons – equivalent to menus – which give quick access to the more frequently used functions, including to features which are specific to DGT-OmegaT, namely: DGT applications (**DocFinder**, **Quest**, **Euramis** and **IATE**), **Revision Mode** and **View Other Target Languages**.



In the FOSS version of DGT-OmegaT, the **DocFinder**, **Euramis** and **Quest** icons are greyed as these databases cannot be accessed outside DGT. The link to **IATE** is active and gives access to its publicly available version⁽⁹⁾.

Although **Euramis** is not available outside EU institutions, most of the EU legislation has been made freely available on the Internet in aligned files⁽¹⁰⁾.

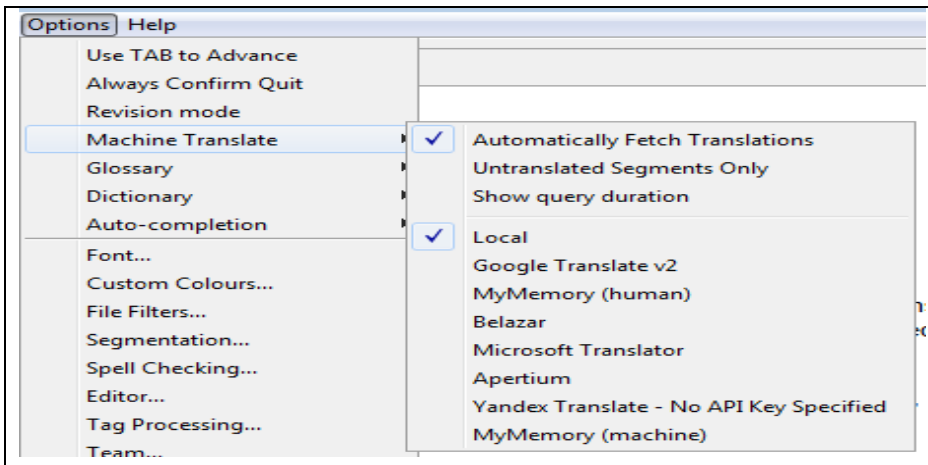
5.2. View Source and Target Files

While working on a translation, the source file of the active document in the **Editor** can be opened directly from DGT-OmegaT by selecting **View source file** in the **Project** menu. This will launch the native application associated with the source document's file type and displays the source content.

There is no (real-time) preview of the translated document in OmegaT, but the completely or partially translated file that is active in the **Editor** can be opened from DGT-OmegaT by selecting **View target file** in the **Project** menu. This will launch the native application associated with the source document's file type and displays the translated content. But there is no interaction between DGT-OmegaT and the native application, so no changes should be made to the target file in the native application as they will not be returned to OmegaT!

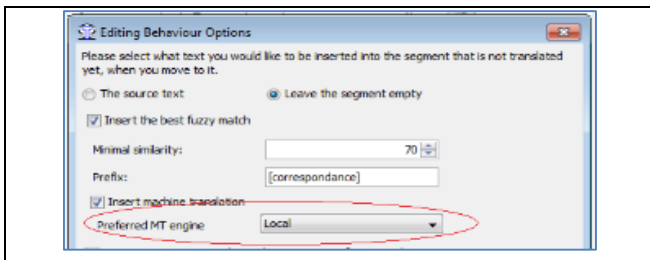
5.3. Machine Translation

DGT doesn't use any of the public MT systems available for several reasons, notably confidentiality and copyright. So a **Local** option has been developed for DGT-OmegaT, adding another implementation of Machine Translation. Instead of calling a server in real-time, this **Local** MT implementation reads data from one or more TMX files. In DGT, only the **Local** MT engine is available.

	<p>In the DGT-OmegaT FOSS version, all engines available in OmegaT 3.6 are in a plugin and will be displayed in the menu only if the plugin is installed. In the executable version we have published, this plugin is installed.</p>
---	--

By creating a new project subfolder named `\mt` and copying the MT file(s) there – something which in DGT is automatically done by the Wizard when creating/updating a project – MT output is displayed in the **Machine Translation** pane, not in the **Fuzzy Matches** pane.

Both with and without the plugin, the **Editor Behaviour Options** window now allows for the automatic insertion of MT if there is no Fuzzy Match above the defined minimal similarity threshold from the project memory or from the external memories.

	<p>If more than one engine is available (if any MT plugin is present), the window also includes the possibility of selecting which engine will be used for automatic insertion.</p>
---	---

5.4. Colours and identification of suggested text in the open segment in the Editor

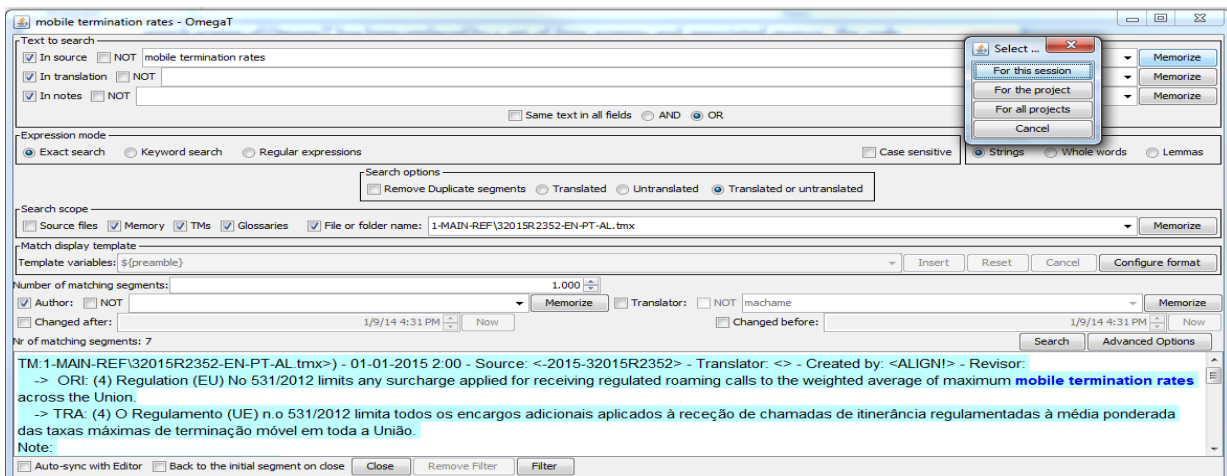
As OmegaT implements a lot of automatic insertion mechanisms, when a new segment is opened it is very important to know if the suggested translation in the editing zone was already in the project memory or if it has been automatically inserted, but not validated. OmegaT only adds a background colour if the segment comes from TMX files in the **tm/auto** or **tm/mt** subfolders. DGT-OmegaT adds colours in some other cases and also displays more information in the starting segment delimiter.

<p>Empty segment</p> <p>36 <segment 0440 +1 more ** EMPTY ** > <end segment></p>	<p>Translated segment</p> <p>Translation last modified by cordoth on 16-mai-2017 Brussels, <segment 0001 ** ALREADY TRANSLATED ** > Bruzelas <end segment></p>	<p>Auto-populated segment</p> <p>Last modified by unknown All of these challenges require huge <segment 0058 ** TM/AUTO ** > Todos estes desafios exigem enormes <end segment></p>
<p>100% match</p> <p>GENERAL PROVISIONS <segment 0426 Match 100/100/100%> DISPOSIÇÕES GERAIS <end segment></p>	<p>Partial match</p> <p>Furthermore, the Commission will: <segment 0063 +4 more Match 100/75/66% > Além disso, a Comissão: <end segment></p>	<p>Machine Translation</p> <p>See Communication 2014/C 188/02. <segment 0179 MT > Ver Comunicação 2014/C 188/02. <end segment></p>

5.5. Search windows: Search Project, Search Directory, Search and Replace and Search and Pre-translate

The OmegaT search feature is already very sophisticated. However, in DGT-OmegaT, the single Search window of OmegaT has been replaced by a set of four windows and associated menus which still look similar to the public OmegaT, but the code has been completely rewritten.

Thanks to this division, the windows are less saturated... and more options could be added. These features are really worth exploring!

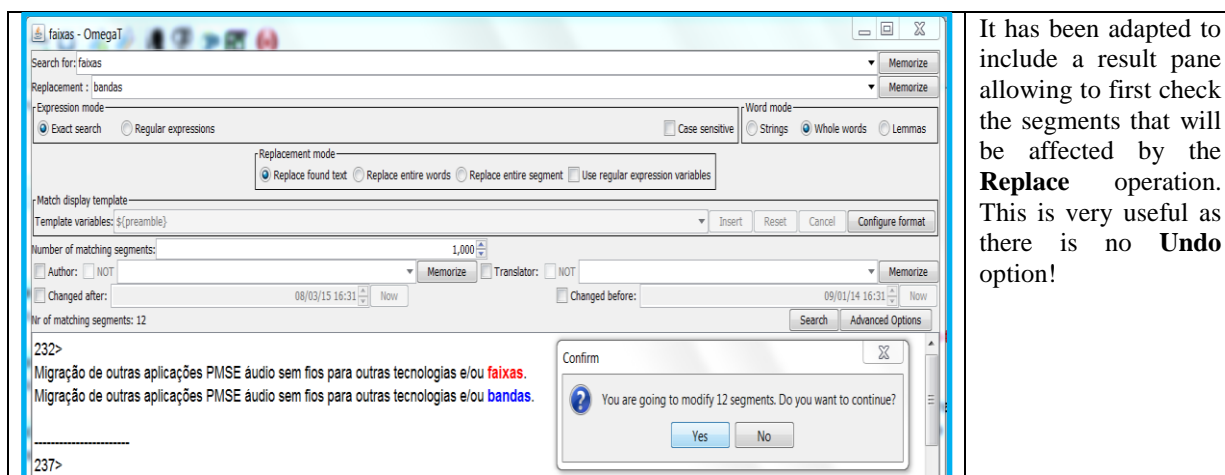


Some of the new options in the **Search Project** window are:

- **Expression mode** and **Word mode** which, combined, give a wider range of options and results.
- **Word mode.** Search options in OmegaT are string-based: when searching for "test", if the segment contains "protestation", it will be found. DGT-OmegaT adds two alternatives:
 - **Whole words:** As in many other edition tools, the previous sample will be rejected unless the search explicitly accepts characters before and after by using the wildcard «?» (one character) or «*» (0 or more characters)
 - **Lemmas:** In this mode, the screen will use tokenizers exactly as for the **Fuzzy Matches** pane: grammatical inflexions of a word will be recognized, while words "containing" another one will not. However, this is still a partial search (i.e. searching for a segment containing the given words) without calculation of a score, while for the **Fuzzy Matches** pane a full search is made and a score calculated.

- **Search in** has been expanded allowing to easily use the **Booleans OR, AND** or/and **NOT**, a very useful feature for translators for terminology purposes. In particular it allows to easily check whether a term/expression has been consistently translated both in the project and in the external memories.
- **Search by file or folder name**, allowing to limit the search to one memory (TMX) file or to a subfolder with several TMX files in the **\tm** folder.
- **Memorize**: Unlike in OmegaT, memorization of a search is done via a button (to be able to selectively memorize only relevant searches) and there are options to memorize the search for any project, the current project or only that session. Furthermore some frequently used Regular Expressions are memorized by default.
- **Match display template**: Display is configurable (as for **Fuzzy Matches**) using the **Config format option**.

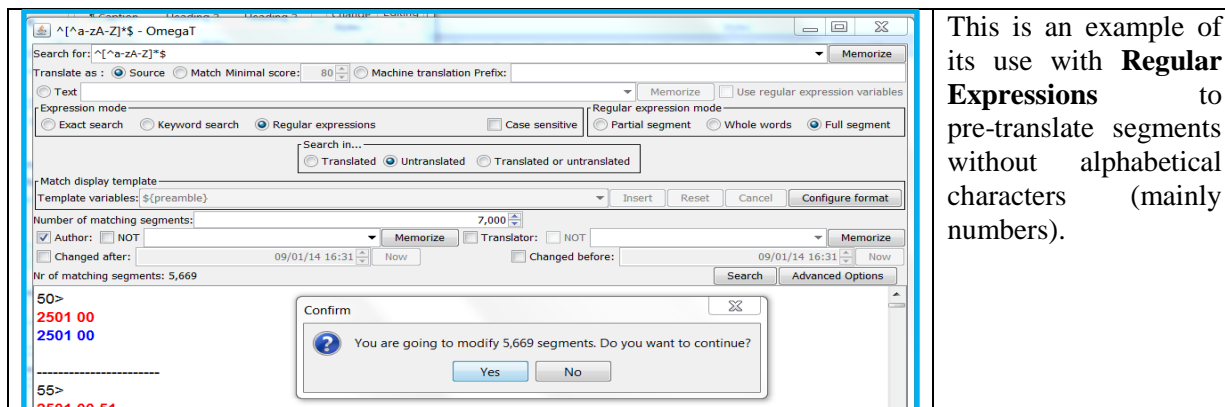
The **Search and Replace** window follows the same approach as the **Search Project** window.



It has been adapted to include a result pane allowing to first check the segments that will be affected by the **Replace** operation. This is very useful as there is no **Undo** option!

The **Search and Pre-translate** window is specific to DGT-OmegaT. Its role is to quickly fill segments for which the translation is either trivial (for example, numbers and other non-translatable characters) or is best done in a batch operation. It works in a similar way as **Search and Replace**, but it does the search in the **source** segment. So it can be used for untranslated segments.

While the **Search and Replace** window only allows to translate with fixed text, **Pre-Translation** is possible by filling the translation with: **Source**: copying source to target; **Match**: pre-translating from external memories matches (match threshold can be defined) and **Text**: as for **Search and Replace**, a constant string with the possibility to use variables recognized during the search.

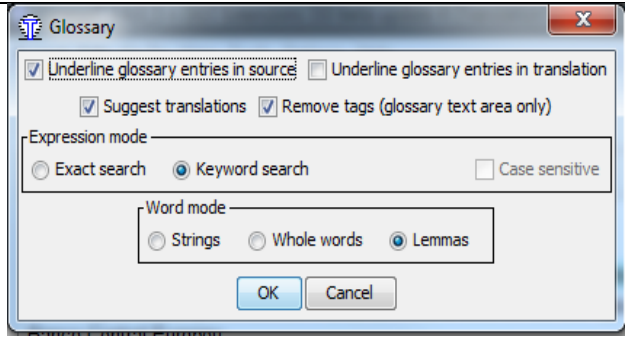


This is an example of its use with **Regular Expressions** to pre-translate segments without alphabetical characters (mainly numbers).

The **Search Directory** follows a similar approach, with adaptations for its particular purpose.

5.6. Glossary

The **Glossary** pane uses colours to differentiate source term, target term and 3rd field (see screenshot in point 5) and the last field is displayed together with the respective target term. Furthermore, the entries after the writable glossary entries are displayed in alphabetical order making it more user-friendly.

	<p>The submenu Glossary - Glossary Pane and Transtips configuration gives the translator some important options concerning terminology display and synchronizes the word search algorithms in order to have consistent results in both displays.</p>
---	---

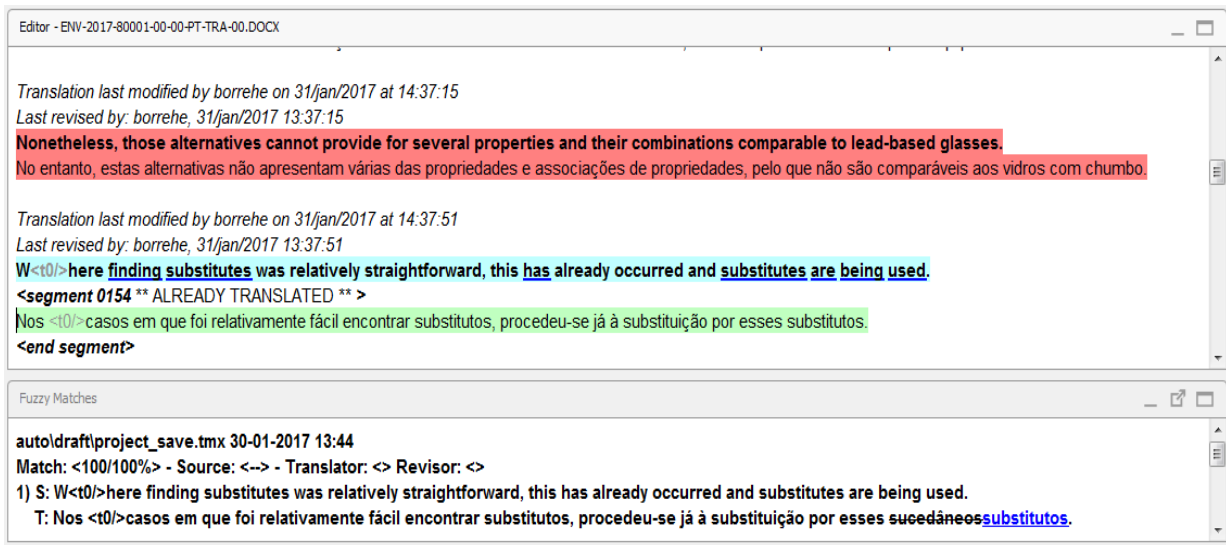
5.7. Match statistics (per file) as a background operation


Match Statistics and **Match Statistics per File** are done in the background thereby allowing the translator to start translating while those statistics are being generated. This feature is especially interesting for large projects of hundreds/thousands of pages with large external translation memories as generating **Match Statistics (per File)** in those projects takes quite some time.

5.8. Revision mode

In DGT, translators can also act as revisers and generally the translator has the last word, which means that, after a document is revised by a fellow translator acting as a reviser, the translator will check the changes made and can accept or reject them. Bearing that in mind, a revision workflow was developed in the DGT-OmegaT Wizard and the following features were developed in DGT-OmegaT.

In DGT-OmegaT, there are now 2 modes - **Translation Mode** and **Revision Mode** – and segments have 3 statuses: untranslated, translated and revised.



When the reviser selects the **Revision Mode** :

- Segments opened and in any way saved in the **Editor** — either changed or not — are marked with **Last revised by** and the login of the reviser. This information is displayed for those segments in the **Editor** and, if so configured, in the **Fuzzy Matches** and **Search** windows. This mark cannot be deleted... but an export can be made without that mark.
- By having the unrevised project memory copied to the `\tm` folder (identified as Draft, for instance), track changes are displayed in target as long as the new option **View diff in target** is activated in the **External TMX** options.
- There is the new option **Go to Next Unrevised Segment** with the shortcut **Ctrl+U** (the same used in the **Translation Mode** to **Go to Next Untranslated Segment**).

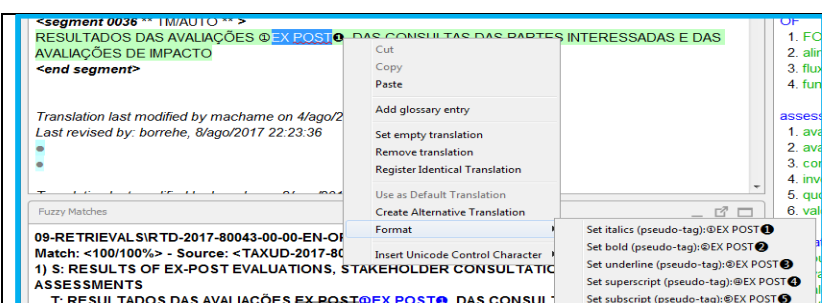
When the translator receives the revised project to finalize, (s)he can easily and quickly check the segments changed by the reviser with these 2 new options:

- **Mark Revised and Changed Segments** in the **View** menu, whereby the segments changed by the reviser are displayed with a red background.
- **Go To Next/Previous Revised & Changed Segment** (and its shortcuts) thereby successively opening only the changed segments for checking.

The **Revision Mode** can also be used by the translator when a particular project is not revised by a colleague, in which case the translator is his/her own reviser as happens in a substantial number of documents.

5.9. Pseudo-tags: insertion of tags in target not present in source

In DGT-OmegaT, there is also an option to add formatting to the target which does not exist in the source segment using the feature **Format (pseudo-tags)**.




Highlighting the text to be formatted, right-clicking the mouse and selecting the desired format (italics, bold, underline, superscript and subscript) inserts characters which exist in Unicode but are rarely used.

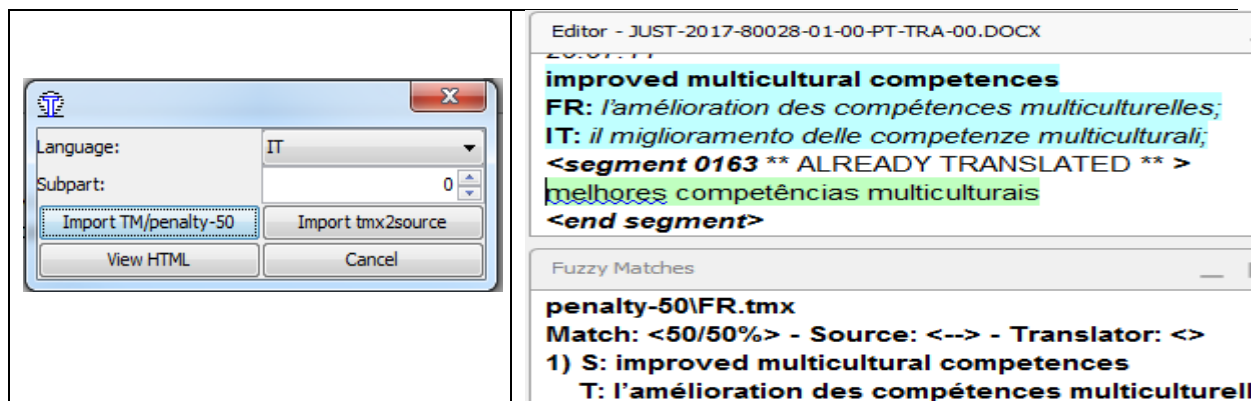
When the translated documents are generated, the **Reformatter** script transforms those pseudo-tags in the corresponding target format. However, pseudo-tags should be used sparingly as sometimes other formatting in the segment may be affected. It is not perfect (and it works only for DOCX files, for the moment), but it can be very useful!

5.10. View other target languages (Cross-lingual concordance)

A substantial part of EU documents is translated simultaneously into many/all EU official languages. So for the purpose of multilingual consistency, it is useful that translators can have a look at ongoing translations into other languages. Omega-T already has a good feature for that purpose: the **tmx2source** folder – but the problem is that files have to be requested and copied manually to it.

For that purpose, the installation of SDL Studio in DGT has the following adaptation: each time a translation is saved (as ongoing, not final!) the SDLXLIFF file is copied to a common folder. At any time, the translator of the same document in another language can use a contextual menu (plugin inside Studio) which will convert this SDLXLIFF file into HTML and display it in the browser.

In DGT-OmegaT, **View Other Target Languages** —  — combines this with other options. First it works in both directions – OmegaT users can see Studio documents and vice versa. Second, DGT-OmegaT users can choose to have them automatically converted to TMX (not to HTML) and copied to the **tmx2source** folder (for display in the **Editor**) and/or to the **tm/penalty-50** folder (for display in the **Fuzzy Matches** pane).



To be continued in Part 2 ...

Maria.Machado@ec.europa.eu
Thomas.Cordonnier@ext.ec.europa.eu
Hilario.Fontes@ec.europa.eu

-
- (1) Machado, M. J., 'Free and open-source software – a translator's good friend', *a folha*, no. 45 — Summer 2014, http://ec.europa.eu/translation/portuguese/magazine/documents/folha45_foss_en.pdf.
- (2) OmegaT, <http://omegat.org/>.
- (3) OmegaT, Download, <http://omegat.org/download>.
- (4) European Commission, *7th RTD Framework Programme* (http://ec.europa.eu/research/fp7/index_en.cfm) and *Horizon 2020* (<https://ec.europa.eu/programmes/horizon2020/>).
- (5) European Commission, *Open source software strategy*, http://ec.europa.eu/dgs/informatics/oss_tech/index_en.htm.
- (6) European Commission: Directorate-General for Translation, *Translation tools and workflow*, <https://publications.europa.eu/en/publication-detail/-/publication/00e51a8e-9c50-11e6-868c-01aa75ed71a1>.
- (7) European Commission, *Machine translation for public administrations — MT@EC*, https://ec.europa.eu/info/resources-partners/machine-translation-public-administrations-mtec_en.
- (8) Moses, <http://www.statmt.org/moses/>.
- (9) Inter-Active Terminology for Europe (IATE), <http://iate.europa.eu/>.
- (10) Fontes, H. L., 'European Union multilingual corpora for reuse in translation', *a folha*, no. 45 — Summer 2014: http://ec.europa.eu/translation/portuguese/magazine/documents/folha45_corpora_en.pdf.