Use of MT@EC by translators in the European Commission

Katja Mai
Translator DGT D.DE.2
Member of the MT User Group

2nd ELRC Conference
Brussels, 26/10 2016
What is MT@EC?

Technology:

- MT system introduced in 2010 in DGT
- Real-life trial with DGT translators since 2011, released for production in June 2013
- Core: statistical MT based on open-source engine Moses, improved by rules-based pre- and post-processing
- Data: Data basis Euramis with ca. 1 billion segments in 24 languages

Language pairs:

- EN <> all official EU languages
- Total of 552 language pairs (76 LPs directly)
“Maturity Check”

64 translators from 22 language departments contributed more than 17000 binary quality judgements on the usefulness of segment translations (EN → X) across a parallel set of documents.

MT user group analysed and clustered the results

Based on this, 10 LDs opted for automatic MT in the real-life trial
Who can use MT@EC?

... staff of the EU institutions

... public administrations in the EU Member States (CEF)

... EMT*-universities (cooperation projects)

* European Masters in Translation
How is MT@EC used in the Commission?

- Since 2013: available via Web Interface
- Translation of documents and "snippets"
- Since end of 2015: As Plugin for Word and MS Outlook
- Fast access to translation of informations in other languages (gisting)
Translation of "snippets": example DE-EN

Text to translate:

Die Kommission legt heute ihren ersten Bericht über die Fortschritte bei der Umsetzung des Partnerschaftsrahmens für die Zusammenarbeit mit Drittländern vor.

Translate from...

DE  Autodetect

Translate into...

- BG
- CS
- EL
- EN
- ET
- FI

Translated text:

The Commission presented today the first report on the progress of implementation of the partnership framework which provides the framework for cooperation with third countries.
How is MT@EC used by translators in DGT?

- As a resource within the CAT tool: adding of MT-tmx optional
  - Automatic creation of MT-tmx for some LPs
  - Manual requests of MT-tmx via Web-Interface for the other LPs
- MT-segments on displayed in CAT tool as "AT" resource
Display of MT-segments in CAT-Tool

With this amendment No.5 and in line with Council Decision (EU) 2016/1337 of 2 August 2016 concerning the allocation of funds decommitted from projects under the 10th European Development Fund (EDF) for the purpose of replenishing the APF, a further amount of EUR 150 000 000 will be allocated to the Action Programme from decommitted funds of the 10th EDF to mainly support African-led Peace Support Operations in 2016.

It supports the Africa-EU cooperation on peace and security under the Joint Africa-EU Strategy which currently seeks to (i) enhance the political dialogue on peace and security, (ii) operationalise the African Peace and Security Architecture (APSA), including the African Standby Force (ASF), and (iii) provide predictable funding for African-led peace operations.

As a result, the budget table in chapter 4.8 would therefore be revised as follows:

<table>
<thead>
<tr>
<th>EUR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Component in EUR</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>EUR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Component in Euro</td>
</tr>
</tbody>
</table>
The partnership framework provides a forward-looking European approach aimed at stepping up cooperation with countries of origin, transit and destination.

The NIF supports projects prepared by the eligible European financial institutions in the neighbourhood in line with the objectives set out in the strategic guidelines.

... and some typical MT errors

<table>
<thead>
<tr>
<th>English</th>
<th>German</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Federal Ministry of Health finances directly the stakeholders.</td>
<td>Das Bundesministerium der Finanzen unmittelbar Beteiligten.</td>
</tr>
<tr>
<td>He teaches <strong>rap</strong> to Roma children.</td>
<td>Er unterrichtet Anwendungsbestimmungen zur Haushaltsordnung Roma-Kinder.</td>
</tr>
<tr>
<td>This Convention was <strong>adopted</strong> by the International Maritime Organisation (IMO) and <strong>ratified</strong> by all EU Member States.</td>
<td>Dieses Übereinkommen wurde von der Internationalen Seeschifffahrtsorganisation (IMO) und von allen Mitgliedstaaten der Europäischen Union <strong>ratifiziert</strong>.</td>
</tr>
</tbody>
</table>
Developments since 2013

- Usage and feed-back from translators are one of the main sources of inspiration for improvements of MT quality

- Explicit feedback (Language weeks) + implicit feedback (final translations)

- Quality of the MT-Engines has improved from generation to generation (more data, inclusion of rules, formatting, placeholders)
Experiences with MT@EC in the German Language Department | 1

Advantages and added value

- Lexical inspiration
- Offers usable phrases /sentence parts
- Less typing needed
- Speeds up translation work
- Good terminology in technical texts
Experiences with MT@EC in the German Language Department | 2

Disadvantages and remarks:

- Shift from translation to Post-Editing-activities
- Editing MT output requires a different kind of attention than translation or revising HT (declinations and verb forms, completeness, negations, additions/omissions)
- Higher need for revision
- Usability of MT-Output varies according to LP (morphology, syntax) and quality of the original
When do we use MT ?

Text types:

+ Press releases, reports, general communications

- Legal acts and in general texts which require a high degree of legal/conceptual accuracy

Selective use of MT as a resource in the translation process alongside with other language resources is becoming more and more part of the daily work of translators in DGT– in varying degrees depending on text types and the language pairs.
Questions?

katja.mai@ec.europa.eu

Additional infos:
http://ec.europa.eu/dgs/translation/translationresources/machine_translation/