A Tale of Eight Countries or the EU Council Presidency Translator in Retrospect

Kristine Metuzale
Client Relationship Manager
Tilde

Alexandra Soska
Language Services Division
Federal Ministry of the Interior, Building and Community, Germany

Mārcis Pinnis
Chief AI Officer
Tilde
THE PRESIDENCY OF
THE COUNCIL OF THE EU

☑️ The Council of the EU is the key
decision maker in the EU

☑️ Presidency rotation every 6
months
Challenges

- Massive amounts of official documentation & texts
- Fast communication of the Presidency work & activities
- Collaboration between delegates, journalists and Presidency staff
- Facilitation of multilingual communication in Europe
- Challenging languages
Solution
The EU Council Presidency Translator

- Secure cloud-based pan-European MT infrastructure
- Customized NMT solution to the Presidency requirements
- Integration of relevant in-domain data & terminology
- Facilitation of multilingual communication in Europe
- Workspace integration of NMT systems
Powers communication across languages benefiting:

- Presidency staff & delegates
- Professional translators
- Journalists
- Presidency visitors
- All EU citizens
Functionality

- Online translation workspace
- Text snippet, document and website translation
- Terminology lookup
- CAT tool plug-in for SDL Trados Studio
- Website translation widget
- Integrated in https://eu2020.de
- Online CAT environment for non-professional translators
- Access to custom and third party MT systems
Architecture

- Cloud-based solution
- Three components:
  - MT systems
    - Custom systems deployed on Tilde MT
    - Third party systems integrated through interfaces
  - Translation broker
  - Translation interfaces for users
Machine translation systems

- **Custom MT systems**
  - For main translation directions of each presidency
  - Keep up with latest NMT technology
  - Adapted using in-domain data
  - Address country-specific dialects
    - FI↔SV for the Finnish Presidency
    - DE↔EN for the Austrian Presidency

- **Generic MT systems**
  - Cover many language pairs
  - Provided by:
    - eTranslation (for all Presidencies)
    - DeepL (for the German Presidency)

<table>
<thead>
<tr>
<th>NMT Toolkit</th>
<th>NMT Architecture</th>
<th>Language pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nematus/AmuNMT</td>
<td>MLSTM</td>
<td>EN↔ET (till 2018), EN↔BG, EN↔DE (for Austrian Presidency)</td>
</tr>
<tr>
<td>Sockeye</td>
<td>Transformer</td>
<td>EN↔ET (since 2018)</td>
</tr>
<tr>
<td>Marian</td>
<td>Transformer</td>
<td>EN↔DE (since 2020), EN↔RO, EN↔FI, FI↔SV, ET↔FI, EN↔HR, DE↔IT, DE↔ES, DE↔PL, DE↔FR</td>
</tr>
</tbody>
</table>


Usage analysis – overall statistics

• Total amount* of translations:
  • **157.7 million words**
  • 14.9 million sentences
  • 4.9 million requests

~300-400 years of work for a productive translator

~1017 books of the size of an average Harry Potter book

*Since the official launch of the EU Council Presidency Translator in 2017 till September 30, 2020
Usage analysis – growing trend

• Translation volume is increasing
• The eu2020.de widget contributes strongly
Usage analysis – custom vs. generic systems

- Custom systems process the majority of translations
- The eu2020.de widget has increased the use of generic engines
Usage analysis – custom NMT system use over time

- Users continue using the custom NMT systems long after presidencies conclude
Usage analysis – primary systems of each Presidency

- Usage differs from country to country
- Higher use evident in countries that integrated the Translator in the official Web sites of the Presidency
Usage analysis – primary systems of each Presidency

- Primary system usage during each Presidency allows us to analyse the adoption of different interfaces by each EU Council Presidency

<table>
<thead>
<tr>
<th>Month</th>
<th>Language</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1</td>
<td>EE (2017)</td>
<td>0%</td>
</tr>
<tr>
<td>M2</td>
<td>EE (2017)</td>
<td>20%</td>
</tr>
<tr>
<td>M3</td>
<td>EE (2017)</td>
<td>40%</td>
</tr>
<tr>
<td>M4</td>
<td>EE (2017)</td>
<td>60%</td>
</tr>
<tr>
<td>M5</td>
<td>EE (2017)</td>
<td>80%</td>
</tr>
<tr>
<td>M6</td>
<td>EE (2017)</td>
<td>100%</td>
</tr>
</tbody>
</table>

Note: The diagram shows the usage analysis for different months and languages, indicating the percentage of usage for each system during each Presidency.
Usage analysis – primary systems of each Presidency

- The CAT tool interface was the most popular in Finland (960 thousand words per month translated on average during the presidency)

![Usage Analysis Diagram]
Usage analysis – primary systems of each Presidency

- The CAT tool interface was not used in Bulgaria, and Croatia
- In Romania, translators started using the CAT interface in the last month of the presidency
Usage analysis – primary systems of each Presidency

- Text translation was the most popular in Estonia, Bulgaria, Austria, and Croatia
- In Croatia, an average of 2.5 million words were translated this way each month
Usage analysis – primary systems of each Presidency

- Document translation was the most popular in Romania, however the most words were translated during the Croatian Presidency (2.1 million words on average each month)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>M1</td>
<td>EN↔ET</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M2</td>
<td>EN↔BG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M3</td>
<td>EN↔DE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M4</td>
<td>EN↔RO</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M5</td>
<td>EN↔FI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M6</td>
<td>EN↔HR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Language Systems:
- EN↔DE, DE↔IT, DE↔ES, DE↔PL, DE↔FR
Translators’ perspective

Contribution
- Collecting language resources
- Testing quality

Benefits
- Learning about MT
- Working more productively
- Feeling empowered
Conclusions

- We presented the EU Council Presidency Translator, a secure cloud-based solution that integrates MT systems from different MT providers and implements a wide spectrum of interfaces for end-users.

- We presented its functionality and analysis of its adoption by users in seven EU Council Presidencies.

- Although the use differs from country to country, we see that there is a need for all types of offered translation interfaces.

- The use of the public interfaces depends on how well the EU Council Presidency Translator is exposed to users in the official web sites of the Presidencies.

- The use of CAT tool interfaces depends on each country’s translation practices.
Thank you!

Kristine Metuzalé
Client Relationship Manager
Tilde
kristine.metuzal@tilde.com

Alexandra Soska
Language Services Division,
Federal Ministry of the Interior,
Building and Community, Germany
alexandra.soska@bmi.bund.de

Mārcis Pinnis
Chief AI Officer
Tilde
marcis.pinnis@tilde.com