European Commission

# MT@EC

---

# The Commission's Machine Translation Service

European Commission

# MT@EC

## Decided 2010:

- Continuity
- Security

## Released 26 June 2013 (version 1.0)

*Version 3.0 released on 1 July 2015*

## Technology:

- Moses statistical machine translation

European Commission

# MT@EC

**Access:** - web user interface (human to machine)
- web services (machine to machine)

**Special features:**

- **User interface** in 24 languages
- Source document format/formatting **maintained** *[not for pdf]*
- Specific output **formats for translation:** tmx and xliff
- Translation can be returned by **email**
- Can translate **multiple** documents to **multiple** languages
- Indication of **quality** for language pairs (using BLEU Scores)
- **Feedback** mechanism (using EU Survey)

European Commission

# MT@EC

## Language coverage:

- All combinations of 24 EU official languages
- Direct translation currently available for:
  - **all official languages having EN as source or target language**
  - **French into and from:**
    **BG, CS, DE, EL, EN, ES, HR, HU, IT, NL, PL, PT, RO, SK, SL**

## Input formats accepted:

- Plain text
- doc(x), xls(x), ppt(x), odt, rtf
- html, xml, xliff
- pdf

European Commission

# Built with translations from:

- European Commission
- European Parliament
- European Council
- Committees (EESC, Regions)
- Court of Auditors
- Court of Justice case law

European Commission

# How much data?

- **750 million monolingual segments (Nov. 2015)**
  Growing at 2.6 million per month

- **98 million monolingual.en**
- **60 million monolingual.fr**
  - **The rest between 22 and 40 million**
- **3 million monolingual.ga**
- **3 million monolingual.hr**

# MT@EC is available for...

- staff of **European institutions and bodies:**

| | |
|---|---|
| ▪ Commission<br>▪ Parliament<br>▪ Council<br>▪ Court of Justice<br>▪ Court of Auditors | ▪ Economic and Social Committee<br>▪ Committee of the Regions<br>▪ European Central Bank,<br>▪ European Investment Bank<br>etc. |

- **online services** funded or supported by the EU
- real-life trial and pilot projects with **public administrations** in the EU Member States
  collaboration projects with **EMT* Universities**

* *European Master's in Translation*

European Commission

# Connecting

*As an individual…*

European Commission

# MT@EC - Machine Translation

European Commission

About | Feedback | Help

**Translate documents** | Translate text | My translation requests | english | Logout

**Select at least one source file :** Choose files... **OR** Drag files here.

**Translate from...**

**Translate into...**

**Output format:** Same as source | TMX | XLIFF

☑ E-mail me my translation.

☐ Delete after download.

Translate document

## Selected files

- DG Briefing.docx (34 kb) **100%** Remove

European Commission

# MT@EC - Machine Translation

European Commission

About | Feedback | Help

ISA

Translate documents | **Translate text** | My translation requests

english | Logout

**Text to translate :**

Good morning to everyone!

Maximum length : 4000 characters.
Remaining : 3975 characters.

Clear

**Translate from...**

EN

**Translate into...**

| | | | |
|---|---|---|---|
| BG | CS | DA | DE |
| EL | ES | ET | FI |
| FR | GA | HR | HU |
| IT | LT | LV | MT |
| NB | NL | PL | PT |
| RO | SK | SL | SV |

E-mail me my translation.
Delete after download.

Translate text

**Translated text :**

Buongiorno a tutti!

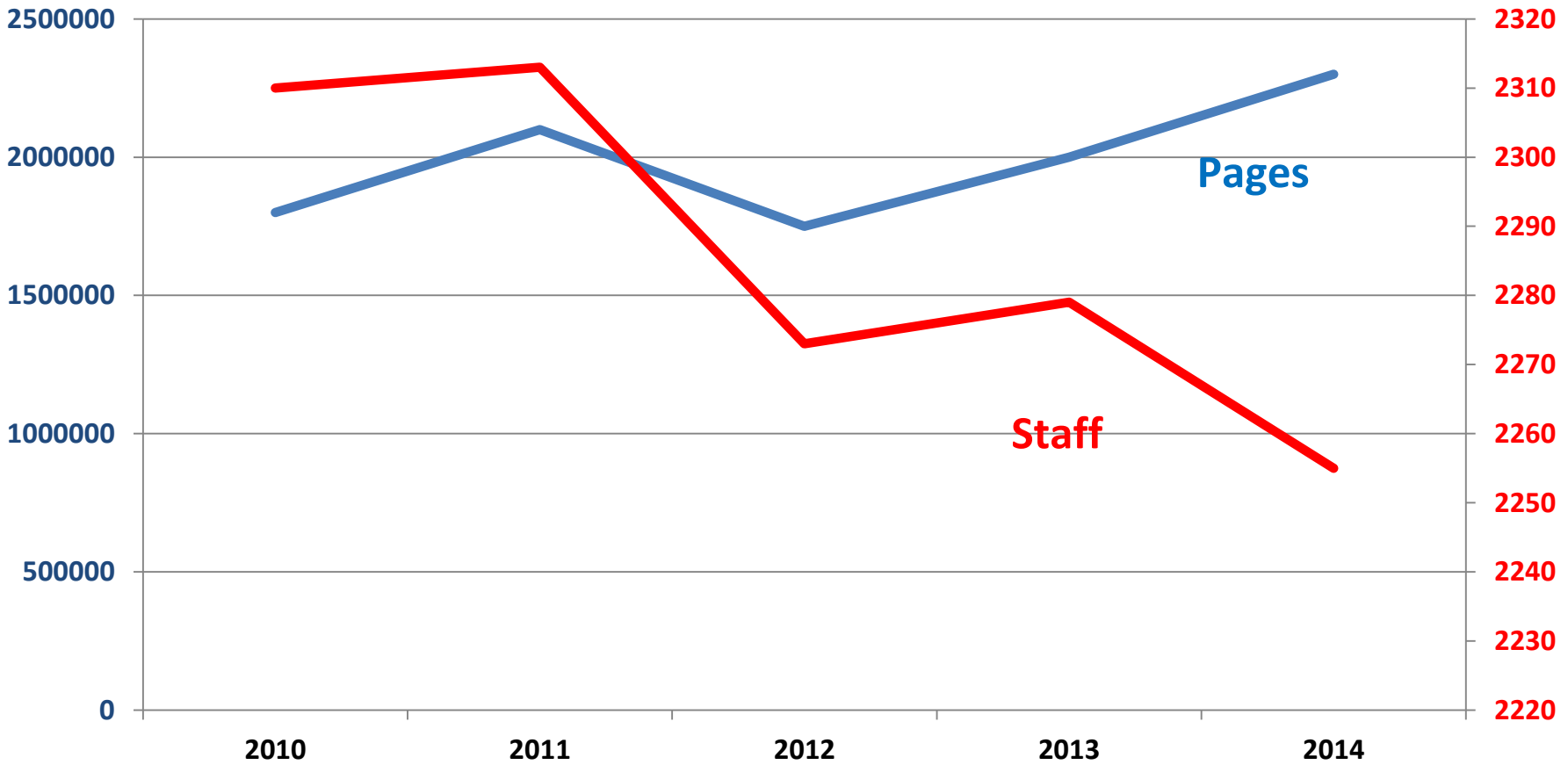
Select all

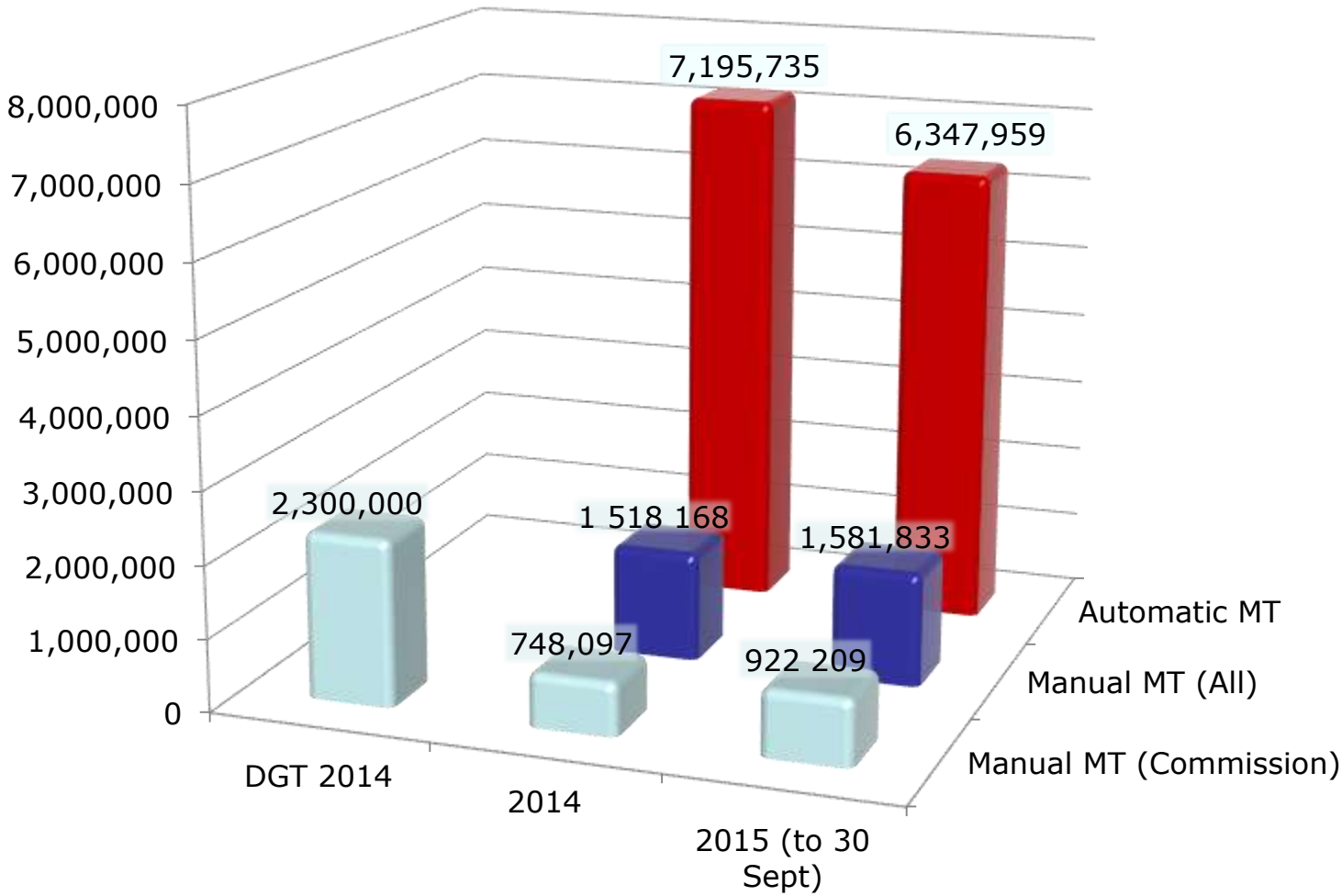
European Commission

# But what do EU translators think?

# Evolution 2010-2014

## Number of translated pages (blue line) and Number of DGT staff (red line)




8,000,000
7,000,000
6,000,000
5,000,000
4,000,000
3,000,000
2,000,000
1,000,000
0
7,195,735
6,347,959
2,300,000
1 518 168
1,581,833
748,097
922 209
DGT 2014
2014
2015 (to 30 Sept)
Automatic MT
Manual MT (All)
Manual MT (Commission)

European Commission

# What's translated automatically?

## Automatic pre-processing for translators:

- Commission
- Council
- Committees

## Online services:

- Internal Market Information System (IMI)
- N-Lex
- Online Dispute Resolution (ODR)
- Pan-European data portal, etc.

European Commission

# Adoption

- *Adopted at language department's request:*
  **Only ET and HU now not included**

- *Most Popular among the:*
  **PT, RO, MT, BG, FR, CS**

- *Least Popular among the:*
  **ET, FI, HU**

# Translator Comments - DGT

+ *Aid for typing*
+ *time savings*
+ *"original" proposed solutions*
+ *guides terminological research*

- *gender/numbers and order of words*
- *can be "fluent", but with mistranslations*
- *omissions and additions*
- *risk of error with incorrect terminology suggested*
- *quality depends on the quality of the originals*

European Commission

# Typographical Errors Survey 2012

- **Paper Cuts:**

  Get feedback from users about the most frequent and irritating – but easy-to-fix – errors

*35 errors fixed, e.g.*

- **Corrupted characters (e.g. in LT: ą, č, ę, ė, į, š, ų, ū, ž)**
- **Compact multiple spaces and commas**
- **Normalise spaces around punctuation (language-specific), e.g. quotation marks, brackets etc.**
- **Add non-breaking space between number and percent**
- **Standardise square and cubic units, e.g. "12 cm³" instead of "12cm3"**

- **Portuguese and Maltese spelling reforms**

European Commission

# 2015: Language Weeks

*Meeting with one language department per week to look at language specific issues*

- **Spurious numbers**
- **Simple rules for set phrases**

# Future Improvements

- *More data*
- *More direct language pairs*
- *Language-specific processing*
- *Domain specific engines*

European Commission

**MT@EC:**

**https://mtatec.ec.testa.eu/mtatec**

**https://webgate.ec.europa.eu/mtatec**

**DGT MT information page:**

**http://ec.europa.eu/dgs/translation/translationresources/machine_translation/index_en.htm**