



# MT@EC

**European Commission machine translation  
for Public Administrations  
in the EU Member States**

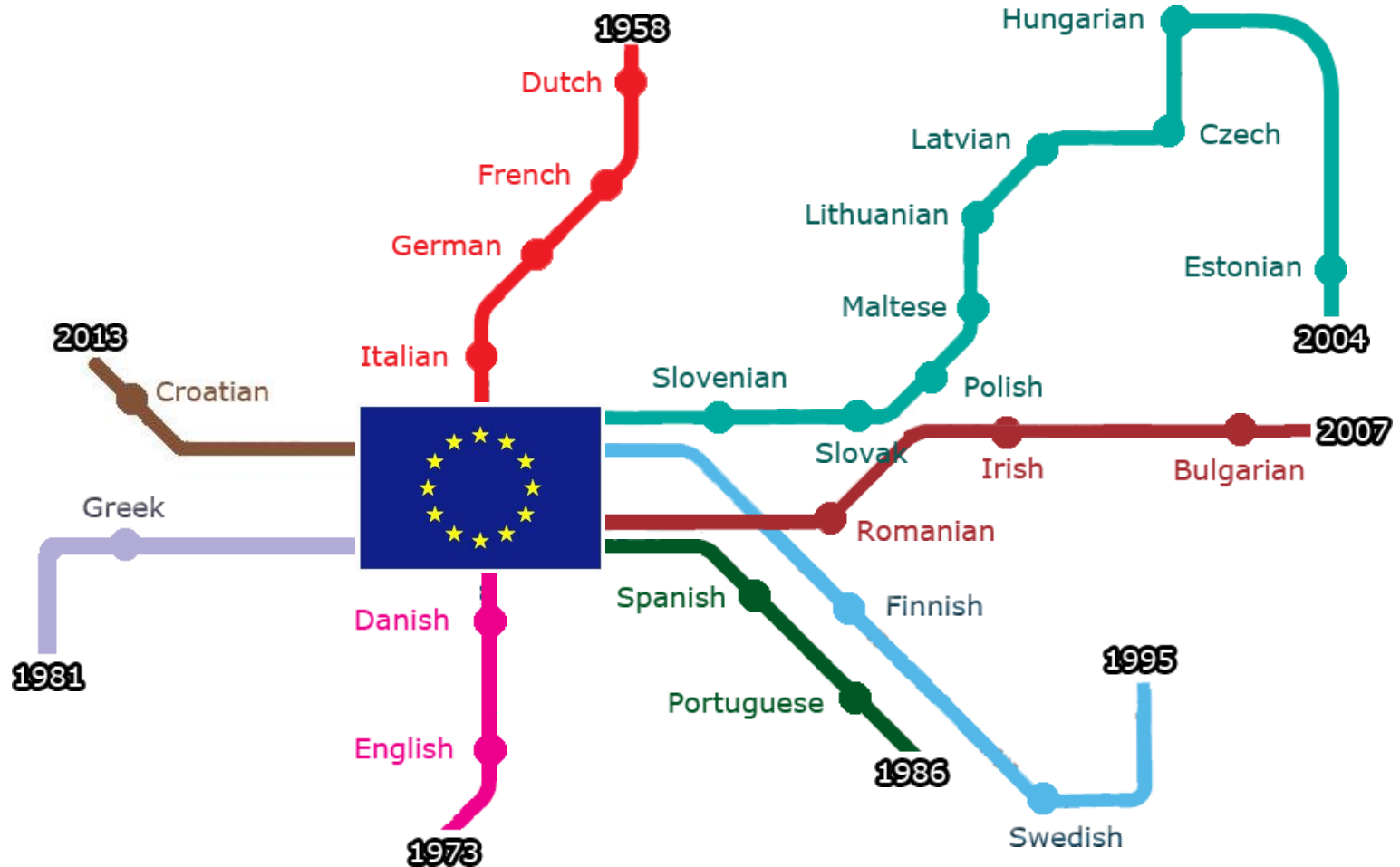
Contact: [Mikael.Reiman@ec.europa.eu](mailto:Mikael.Reiman@ec.europa.eu)

*Updated: 29.10.2014*

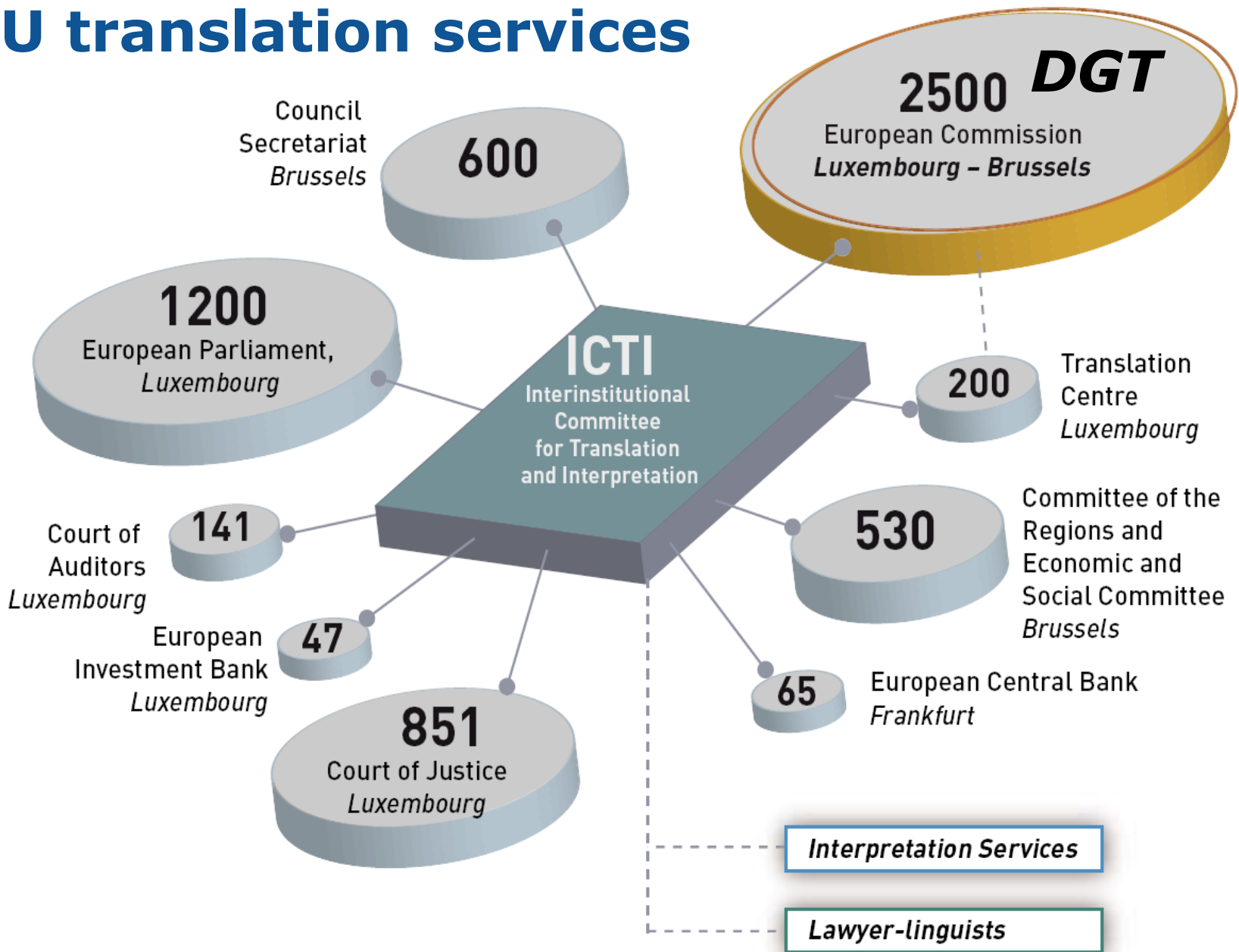
# European Commission machine translation for public administrations

- *European Commission and languages*
- *MT@EC: machine translation for the EU*
- *Access to MT@EC for public administrations*
- *What next?*

# EU official languages over time



# EU translation services



# Why do we need machine translation?

*The Commission...*

- **DGT has 1700 translators**
- **Over 2 M pages translated in 2013**

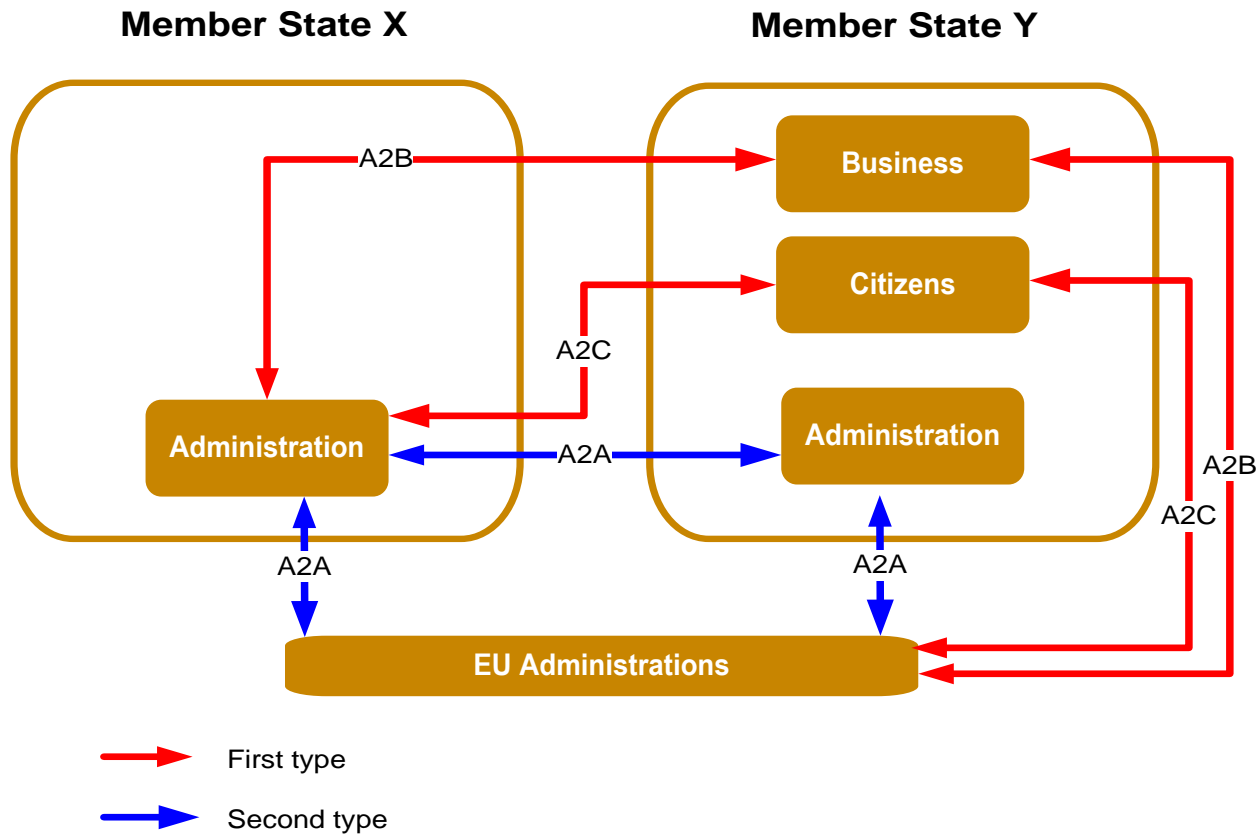
*But...*

**...just to make europa.eu fully multilingual**  
almost 6.8 M documents to be translated  
or 8 500 translators/year!

*The result:*

**Thousands of non-translated documents**  
*(and this does not include user generated content)*

# There are also interactions...



# MT and EC: a long history

## ***Started in the 1970s***

- Eurotra (78-92): research, high expectations
- Rule-based ECMT (75-97), costly to develop – not scalable (18 language pairs in 20 years - coverage of post-2004 languages never attempted- system shut down in 2010)

## ***Data-driven systems (Statistical MT) :***

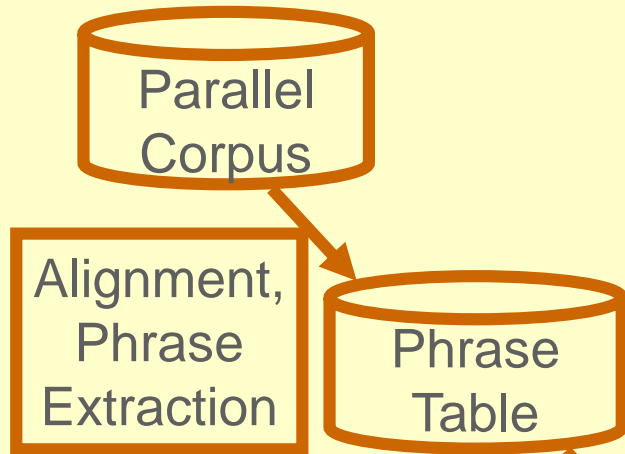
- Economic and quick to develop... if you have good data
- EC needs solution for all EU languages... and has good data

## ***EC action plan (2009), Inter-service task force (2010)***

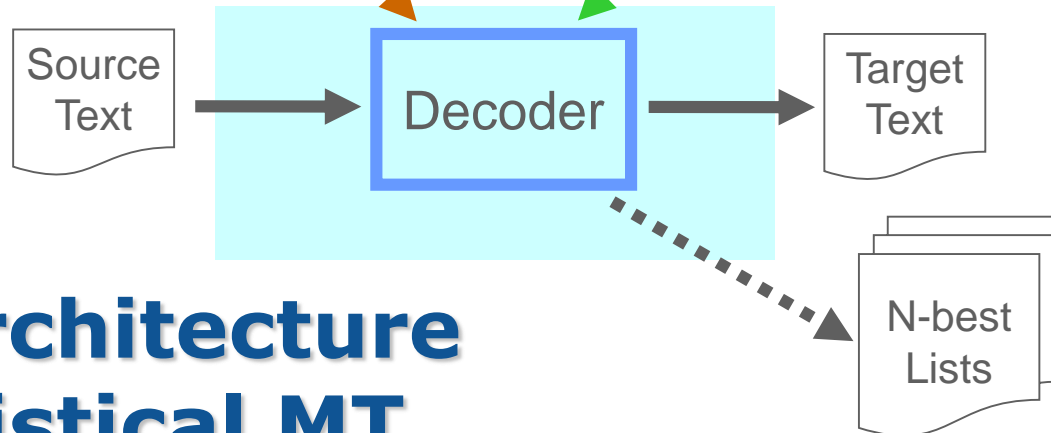
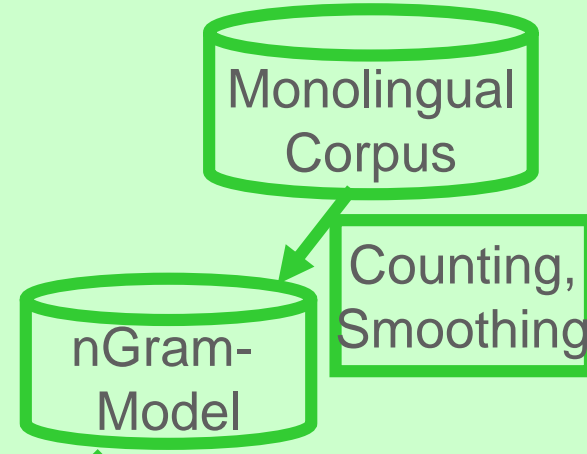
- The goal: MT@EC offering machine translation for all languages to and from English, operational in July 2013



## Translation Model (*Adequacy*)



## Language Model (*Fluency*)

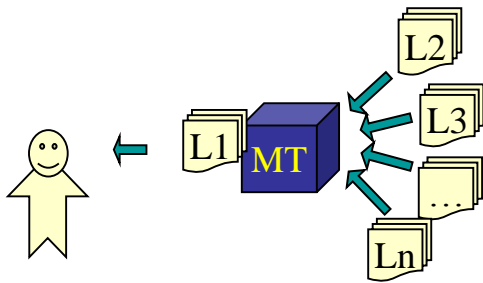


# Basic Architecture for Statistical MT



# Two Usage Scenarios for MT@EC

## *MT for understanding (inbound)*



*Robustness, Coverage*

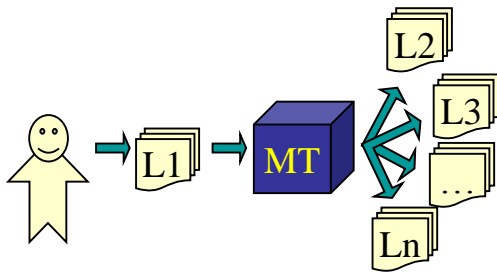
*Practically unlimited demand; free web-based services cover much of it*

## Requirements for MT@EC

- Provide MT as a (simple and robust) service
- Optimise quality for understandability (gisting)
- Deal with many domains, document types, formats, ...
- Scale to huge volumes

# Two Usage Scenarios for MT@EC

## *MT for dissemination (outbound)*



### *Textual quality*

*Publishable quality can only be authored by humans; Translation Memories & CAT-Tools used by professional translators*

## Requirements for MT@EC

- Provide MT as a tool within a CAT workflow
- Develop new ways to incorporate feedback
  - explicit feedback on MT quality, implicit feedback via TM
  - improvements requiring language-specific knowledge
  - towards hybrid approaches
- Optimise quality for post-editing

# MT@EC: a European Commission product

**Released :** 26 June 2013 (version 1.0)

*Version 2.0 released on 3 July 2014*

**Languages:** All 24 EU official languages  
552 language pairs (*62 direct*)

**Technology:** Statistical machine translation  
*using open source software Moses co-funded by EU  
Framework Programmes for research and innovation*

**Development by DGT:** between 2010-2013  
*co-funded by the ISA programme (action 2.8)*

\* Interoperability solutions for public administrations

[http://ec.europa.eu/isa/actions/02-interoperability-architecture/2-8action\\_en.htm](http://ec.europa.eu/isa/actions/02-interoperability-architecture/2-8action_en.htm)

# MT@EC description

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

**Security:** Host (EC data centre) + access (ECAS)  
+ transfer (sTESTA)

## Special features:

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

## **MT@EC is already available to...**

### **... the staff of European institutions and bodies:**

Commission , Parliament, Council, Court of Justice, Court of Auditors, Committees, Central Bank, Investment Bank, Translation Centre, etc.

### **... online services funded or supported by EC**

*Currently:* IMI, SOLVIT

*In the pipeline:* eJustice portal, JoinUp, TED, nLEX, Open Dispute Resolution portal, CircaBC, etc.

# MT@EC for public administrations

Free **real-life trial** until end of **2015**:

- **Staff** can have **direct free access** to the standard MT@EC service (*upon request*)
- Your **Organisation** can participate in a **customisation pilot project**, where DGT can also build specific engines with your data (*bilateral cooperation agreement*)\*

Two levels of **Security** for document transfers possible:

- over sTESTA\*\*, a very secure private network between public administrations in the EU, separate from the internet
- over the internet through a secure (https) connection

\* *Administrative Agreements signed only until end of June 2015*

\*\* *You can check if your organisation has access to sTESTA on:  
<https://portal.testa.eu/jetspeed/portal/homepage/about.psml>*

# Customisation pilots

- Pilot **A**: Connect an information system of yours to the standard MT@EC service.
- Pilot **B**: DGT builds custom engines (your data) available to all through MT@EC
- Pilot **C**: DGT builds custom engines (your data) available only to you through MT@EC
- Pilot **D**: DGT builds custom engines (your data) for you to run in your premises
- Pilot **E**: DGT assists you to build your own custom engines for you to run in your premises

If you are interested  
email [DGT-MT@ec.europa.eu](mailto:DGT-MT@ec.europa.eu)

# Staff access to MT@EC

- Get an individual **ECAS** user name and password (self-registration) using your work email address.  
*[go to <https://webgate.ec.europa.eu/cas/eim/external/register.cgi> and follow the instructions]*
- Send an **email** to [DGT-MT@ec.europa.eu](mailto:DGT-MT@ec.europa.eu) asking for access either over sTESTA or over the internet.
- DGT will activate your access and inform you by email.



# You can ask for documents...



European Commission



english

## Machine Translation

European Commission

Translate documents

Translate text

My translation requests

About Feedback Help Logout

### Translate documents

Select at least one source file :

Choose files...

Translate from...

EN

Translate into...

- |  |                             |                             |                             |
|--|-----------------------------|-----------------------------|-----------------------------|
| <input type="checkbox"/> BG            | <input type="checkbox"/> CS | <input type="checkbox"/> DA | <input type="checkbox"/> DE |
| <input checked="" type="checkbox"/> EL | <input type="checkbox"/> ES | <input type="checkbox"/> ET | <input type="checkbox"/> FI |
| <input type="checkbox"/> FR            | <input type="checkbox"/> GA | <input type="checkbox"/> HR | <input type="checkbox"/> HU |
| <input type="checkbox"/> IT            | <input type="checkbox"/> LT | <input type="checkbox"/> LV | <input type="checkbox"/> MT |
| <input type="checkbox"/> NL            | <input type="checkbox"/> PL | <input type="checkbox"/> PT | <input type="checkbox"/> RO |
| <input type="checkbox"/> SK            | <input type="checkbox"/> SL | <input type="checkbox"/> SV |                             |

Output format:

Same as source  TMX  XLIFF

E-mail me my translation.

Translate document

### Selected files

- MT at the Commission.doc (209 kb) **100%** [Remove](#)
- 000 - Budget flows.xlsx (8 kb) **100%** [Remove](#)

## Tab: Translate documents

- Click on "Choose files..."  
*[notice that only the acceptable formats are proposed]*
- Select the file(s) you wish to translate  
*[notice that these appear in the box "Selected files"]*
- Select language to "Translate from..."  
*[no language detection for the time being...]*
- Select language(s) to "Translate into..."  
*[notice the quality indication (medals)]*
- Select output format  
*["same as source", tmx or xliff]*
- Decide on reception by email

## Tab: Translate documents

- Click on "Submit documents to translate"  
*[Notice message: "Your request has been submitted... "]*
- Wait
- Result (or Notification) sent by email  
*[also available to download under "My Translation requests"]*

... or you can use the text box



## Machine Translation



english

European Commission

Translate documents

Translate text

My translation requests

About Feedback Help Logout

### Translate text

Text to translate :

Machine translation

Maximum length : 4000 characters.  
Remaining : 3981 characters.

Clear

Translate from...

EN

Translate into...

- |                          |                                     |                          |
|--------------------------|-------------------------------------|--------------------------|
| <input type="radio"/> BG | <input type="radio"/> CS            | <input type="radio"/> DA |
| <input type="radio"/> DE | <input type="radio"/> EL            | <input type="radio"/> ES |
| <input type="radio"/> ET | <input type="radio"/> FI            | <input type="radio"/> FR |
| <input type="radio"/> GA | <input checked="" type="radio"/> HR | <input type="radio"/> HU |
| <input type="radio"/> IT | <input type="radio"/> LT            | <input type="radio"/> LV |
| <input type="radio"/> MT | <input type="radio"/> NL            | <input type="radio"/> PL |
| <input type="radio"/> PT | <input type="radio"/> RO            | <input type="radio"/> SK |
| <input type="radio"/> SL | <input type="radio"/> SV            |                          |

E-mail me my translation.

Translate text

Translated text :

Strojno prevodenje

Select all

✓ Your request has been submitted.

## Tab: Translate text

- Type/paste the source text in box "Text to translate" (max: 4000 chars)
- Select language to "Translate from..."  
*[automatic language detection also available]*
- Select language to "Translate into..."
- Decide on reception by email
- Click on "Translate text"
- Wait
- Result appears in box "Translated text"  
*[It may also be sent by email. It is also available for download under "my translation requests"]*

# You can then download the result ...



## Machine Translation



english

European Commission

Translate documents

Translate text

My translation requests

About Feedback Help Logout

### History

Original file name	Status	Message	Document format	Target format	Source language	Target language	Download
000 - Budget flows.xlsx	Ongoing		xlsx		EN	EL	
MT at the Commission.doc	Translated		doc		EN	EL	
Text	Translated		txt	txt	EN	HR	

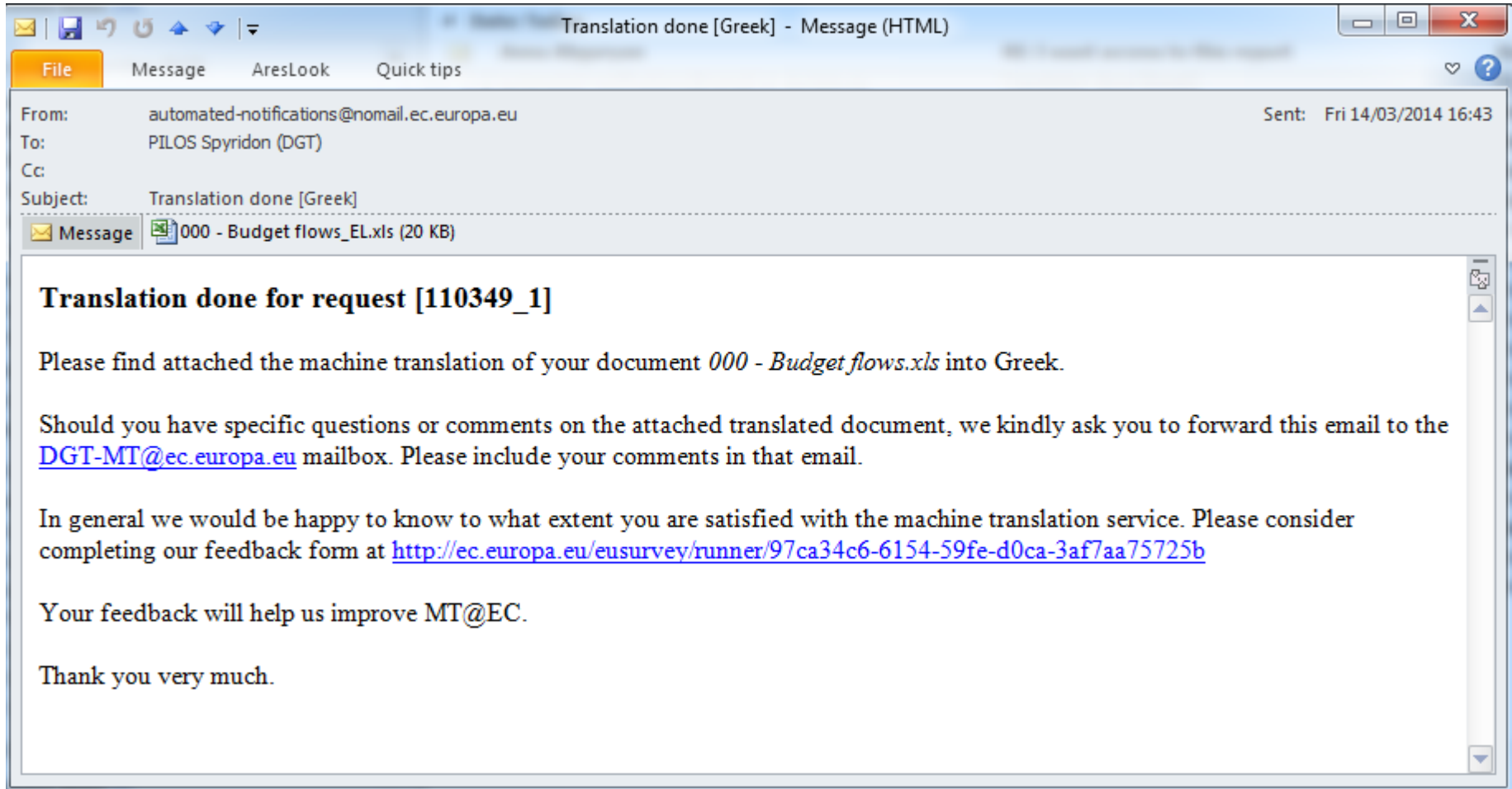
## Tab: My translation requests

- Monitor progress of translation requests
- Links to download each translation (document or text)

### **NOTE:**

**Entries (and download links)  
remain available for 24 h**

... or get it by email





# Email (one per document/request)

The email includes:

- The translated document or text  
*(if email option selected)*
- the notification and link for downloading document  
*(if email delivery option is not selected)*
- Link for providing feedback
- Link to functional mailbox *DGT-MT@ec.europa.eu*



## ***"Maturity Check" (April-May 2011)***

- Can baseline MT engines already be used as such?
- Identify main sources of problems for various languages, cluster them across languages

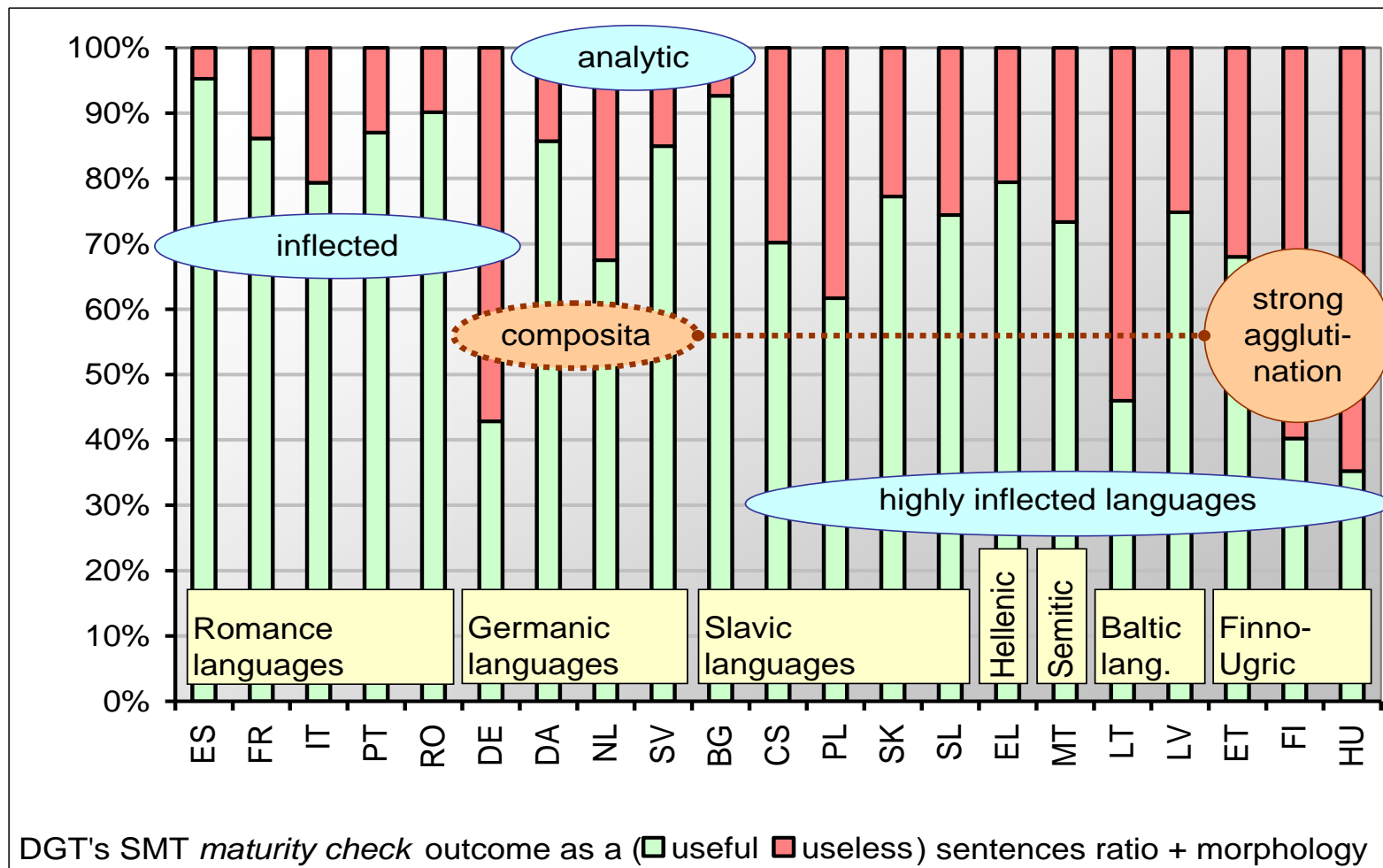
## ***"Real-life trial" (July 2011-June 2013)***

- Make first MT results available to translators
- Auto-MT for 10..19 "best" language pairs from EN (now: all)
- On-demand MT for others

## ***Automatic scores***

- BLEU scores for internal tuning and regression testing.
- Can help to identify domains/document types where MT is most useful, but also point to systematic difficulties.

## Maturity check 2011 (EN->X)



# From the translator's point of view



- + Aid for typing
- + time savings
- + “original” proposed solution
- + guides the terminological research
  - gender/numbers and order of words
  - can be "fluent", but with mistranslations
  - omissions and additions
  - risk of error when incorrect terminology suggested
  - quality dependent on the quality of the originals

# MT@EC: right for the EU

## Quality:

- built on data derived from EU translations  
(*Euramis translation memory system: 800 M segments in 24 languages and annual growth rate > 20%* )
- designed for EU relevant collaboration
- team of **computational linguists** working with **translators and linguists** in DGT
- work to improve MT for **all** EU languages

## Security

## Customer support

# MT@EC: what next

## ***CEF (Connecting Europe Facility)***

- A funding programme for building and deploying infrastructures.
- Includes deploying mature technologies to build, enable and operate pan-European **Digital Services**.
- Includes an Automated Translation (AT) platform as one of its core building blocks for digital services.
- A key component of the AT platform is MT@EC.
- Emphasis is placed on **secure, quality, customisable** machine translation.

## ***MT Conference for Public Administrations***

- on **5 December 2014**, in Brussels
- organised by the European Commission (DG Translation)

# The automated translation platform

- To facilitate cross-border information exchange and enable cross-border access to online content and services provided by the digital service infrastructures of the CEF.
- To offer MT services to EU institutions and public administrations in the Member States.
- To build on the existing Commission Machine Translation service (MT@EC)
- Emphasis is placed on **secure, quality, customisable** machine translation.

→ *Watch this space:*

<http://ec.europa.eu/digital-agenda/en/connecting-europe-facility>



European  
Commission

BLEU	
FI - EN	57,3
EN - FI	45,2
IT - EN	71,0
EN - IT	65,5
MIN	42,3 (EN-HU)
MAX	77,8 (MT-EN)
AVG	62,0



## European Union to boost Ebola research with €24.4 million

The European Commission has today announced €24.4 million from the EU budget for urgently needed Ebola research

The funding will go to five projects, ranging from a large-scale clinical trial of a potential vaccine, to testing existing and novel compounds to treat Ebola

## Euroopan unioni edistää Ebola-viruksen tutkimus 24.4 milj. euroa

Euroopan komissio ilmoitti tänään myöntävänsä 24.4 miljoonaa euroa EU:n talousarviosta tarvitaan pikaisesti Ebola-viruksen tutkimusta

Rahoitusta myönnetään viidelle hankkeelle, jotka vaihtelevat merkittävästi kliinisen lääketutkimuksen mahdollisen, testataan olemassa olevien ja uusien aineiden käsittelemään Ebola-viruksen

**The money from Horizon 2020, the EU research and innovation programme, will be provided via a fast-track procedure in order to start work as soon as possible**

**Rahoitetaan Horisontti 2020 -ohjelmaa, joka on unionin tutkimuksen ja innovoinnin ohjelmassa, tarvitaan nopeutettua menettelyä, jotta voidaan aloittaa mahdollisimman pian**

The Commission is also working with industry on the further development of vaccines, drugs and diagnostics for Ebola and other haemorrhagic diseases within the Innovative Medicines Initiative

Lisäksi komissio tekee yhteistyötä teollisuuden kanssa kehittää edelleen rokotusten, lääkkeiden ja diagnostiikan Ebola-viruksen aiheuttamasta verenvuototaudista yhteisössä ja muut innovatiiviset lääkkeet -aloite

**The President of the European Commission, José Manuel Barroso said: "We're in a race against time on Ebola, and we must address both the emergency situation and at the same time have a long term response"**

I am therefore happy to announce that additional funding of €24.4 million will be made available which will speed up some of the most promising research to develop vaccines and treatments.

**Euroopan komission puheenjohtaja José Manuel Barroso sanoi: "Olemme perustuviin ajan Ebola, ja meidän on otettava huomioon sekä hätätilanteen ja samalla pitkän aikavälin toimia"**

Näin ollen minulla on ilo ilmoittaa, että 24.4 miljoonan euron lisämääräraha asetetaan saataville, joilla nopeutetaan lupaavimpia tutkimusta, jonka tavoitteena on kehittää rokotteita ja hoitoja

**The funding will be provided following proposals from teams across the EU and in numerous 3rd countries, which were evaluated by an independent group of experts**

**Rahoitusta annetaan ryhmien ehdotusten perusteella eu:ssa 3. ja useissa maissa, joissa oli arvioinut riippumaton asiantuntijaryhmä**

The announcement comes as the WHO hosts a high level meeting in Geneva today, involving the international community and industry representatives

Ilmoituksessa on who:lle järjestää korkean tason kokouksessa Genevessä, kansainvälisen yhteisön ja teollisuuden edustajia

The European Union has been active from the early stages of the crisis

Euroopan unioni on jo varhaisesta vaiheesta lähtien kriisiin

**The European Commission alone has so far pledged €180 million of humanitarian and development aid to help the countries affected by the epidemic**

This is done, for instance, by providing immediate healthcare to the affected communities and helping contain the spread of the epidemic through rapid diagnosis and disease awareness campaigns

Besides financial aid, the EU is also contributing to fight the epidemic with experts' presence on the ground and coordination for the delivery of supplies and possible evacuations

**Euroopan komissio on tähän mennessä luvattu 180 miljoonan euron arvosta humanitaarista ja kehitysapua, jolla näitä maita autetaan elämää epidemia vaikeuttaa**

Tämä tapahtuu esimerkiksi säätämällä välittömästi terveydenhuollon vaikutusalueella sijaitseville yhteisöille ja epidemian leviäminen voidaan estää nopeilla ja taudeista tiedottavien kampanjoiden

Lukuun ottamatta eu:lla on myös torjua epidemiaa kanssa asiantuntijoiden läsnäolo kentällä ja koordinointi tavarantoimitus ja mahdollinen evakuointi