



European Commission machine translation collaboration across linguistic barriers

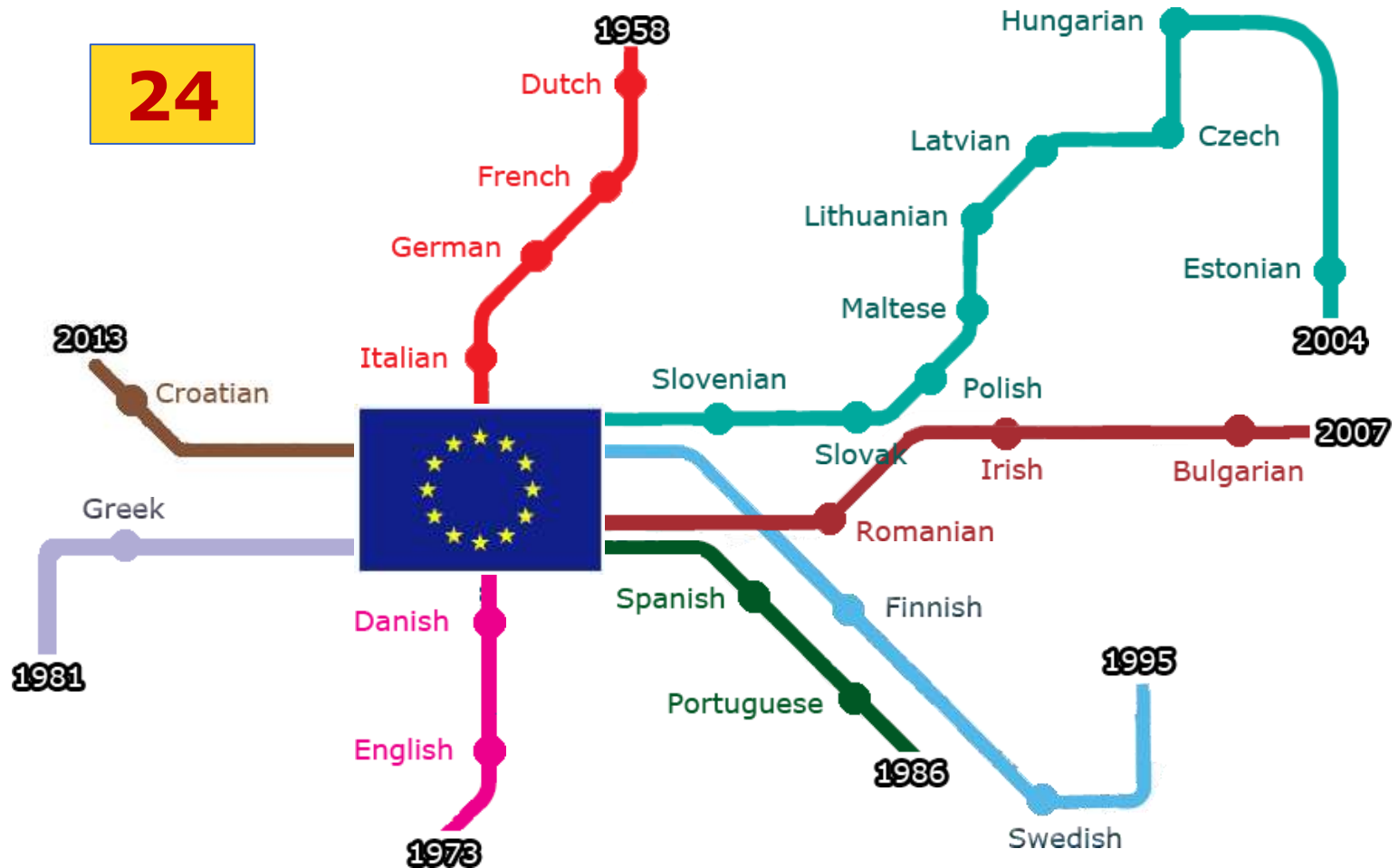
Spyridon Pilos
Head of language applications
Directorate-General for Translation

Riga, 28.04.2015

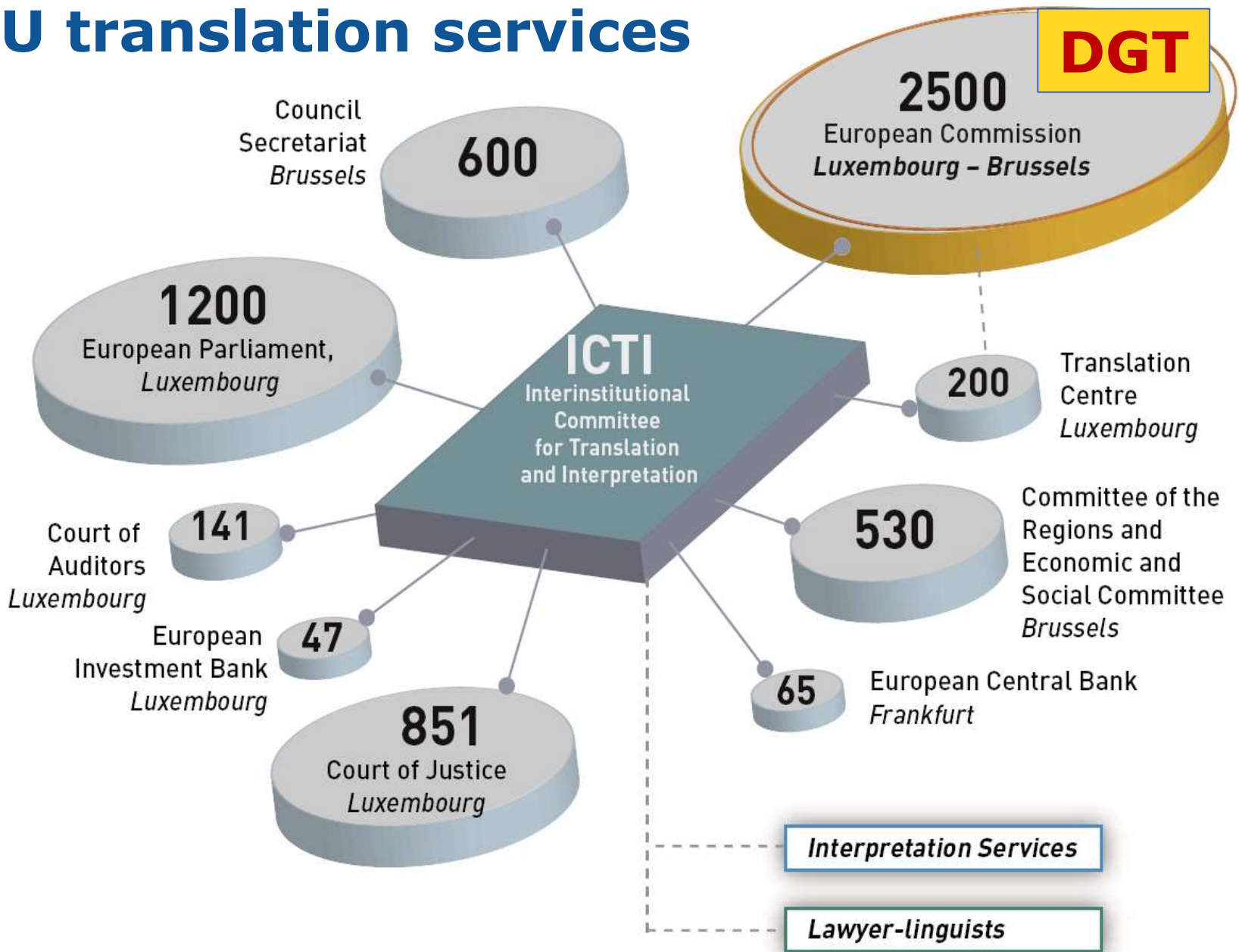
Machine translation for the EU

- *EU and languages*
- *MT@EC and its users*
- *Automated translation platform*

EU official languages over time



EU translation services



Why does the Commission need MT?

The Commission...

- **DGT** has **1700** translators
- **Over 2 M pages** translated in **2014**

But...

...just to make europa.eu fully multilingual

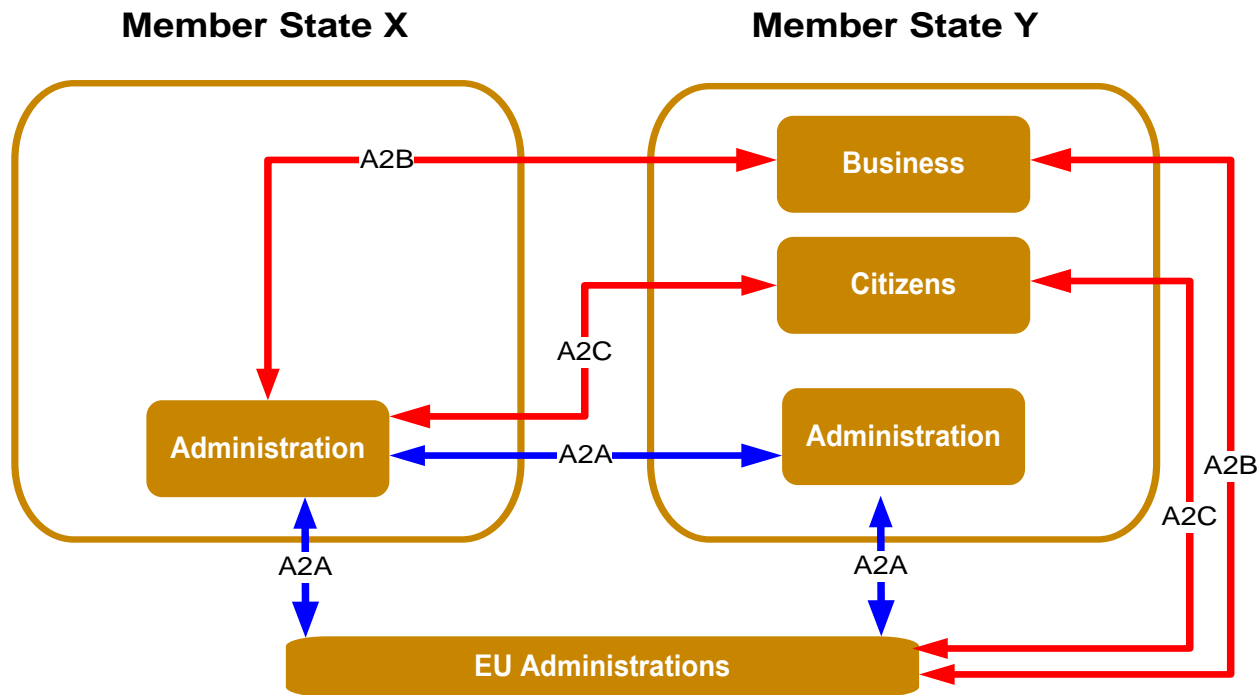
more than **7 M documents** would have to be translated
or **8 500 translators/year!**

The result:

Thousands of non-translated documents

*(and this does not include
user generated content and social media)*

There are also interactions with and between actors in the **Member States**



- First type
- Second type

Vision



Wouldn't it be great if I could start *using a public service*
in any Member State from any place
and obtain the information *in my mother tongue?*





European
Commission



Interoperability
addresses
the need for

Cooperation among public administrations, to establish public services

Exchanging information among public administrations to fulfil legal requirements or political commitments

Sharing and reusing information among public administrations to increase administrative efficiency and cut 'red tape' for citizens and businesses

The
results
are

Improved public service delivery to citizens and businesses, by facilitating the 'one-stop-shop' delivery of public services

Lower costs for public administrations, businesses and citizens due to the efficient delivery of public services



European
Commission



A framework: **EIF**

European **I**nteroperability **F**ramework

http://ec.europa.eu/isa/documents/isa_annex_ii_eif_en.pdf

A programme: **ISA**

Interoperability **S**olutions for public **A**dministrations

<http://ec.europa.eu/isa/>

Sharing results: **Joinup**

An open collaborative platform supporting interoperability in Europe

<http://joinup.ec.europa.eu/>

EIF: 12 principles for EU public services

- Need for EC action** Subsidiarity and Proportionality
- User needs and expectations** User Centricity, Inclusion and Accessibility, Security and Privacy, **Multilingualism**, Administrative Simplification, Transparency, Preservation of Information
- Collaboration** Openness, Reusability, Technological Neutrality and Adaptability, Effectiveness and Efficiency

Recommendation 4. *Public administrations should use information systems and technical architectures that cater for **multilingualism** when establishing a European public service.*

The role of Machine Translation

MT is the only viable solution for:

- quick and cheap **access to information** in foreign languages.
- understanding **information received** in a foreign language that otherwise could not be used or would require substantial time and costs to translate.
- making **multilingual use of websites** possible
- facilitating cross-lingual information **search** and **analytics**.

That is why machine translation (MT) is a **critically important** technology for multilingual Europe





MT@EC: a European Commission product

Released : 26 June 2013

Languages: All directions between
24 EU official languages

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)

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)

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 also 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)



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



[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...

- | | | | |
|---|---|--|--|
| <input type="radio"/> BG  | <input type="radio"/> CS  | <input type="radio"/> DA  | <input type="radio"/> DE  |
| <input type="radio"/> EL  | <input checked="" type="radio"/> ES  | <input type="radio"/> ET  | <input type="radio"/> FI  |
| <input type="radio"/> FR  | <input type="radio"/> GA  | <input type="radio"/> HR  | <input type="radio"/> HU  |
| <input checked="" type="radio"/> IT  | <input type="radio"/> LT  | <input type="radio"/> LV  | <input type="radio"/> MT  |
| <input type="radio"/> NB  | <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.
 Delete after download.

[Translate text](#)

Translated text :

Buongiorno a tutti!

[Select all](#)

Who can use MT@EC today ?

→ *European institutions and bodies:*

- | | |
|--|---|
| <ul style="list-style-type: none">▪ Commission▪ Parliament▪ Council▪ Court of Justice▪ Court of Auditors | <ul style="list-style-type: none">▪ Economic and Social Committee▪ Committee of the Regions▪ European Central Bank,▪ European Investment Bank
etc. |
|--|---|

→ *Online services funded or supported by the EU*

→ *Public administrations in the EU Member States*



Country	reg'd	using
Austria	4	4
Belgium	6	5
Bulgaria	8	8
Croatia	0	0
Cyprus	81	55
Czech Republic	30	25
Denmark	1	0
Estonia	5	4
Finland	3	3
France	34	25
Germany	47	37
Greece	46	34
Hungary	1	0
Ireland	0	0

Users in public administrations

registered
520
using
359
69,4%

If you are wish to try it
request access to
DGT-MT@ec.europa.eu

Country	reg'd	using
Italy	36	24
Latvia	4	3
Lithuania	9	8
Luxembourg	4	3
Malta	0	0
Netherlands	14	14
Poland	8	7
Portugal	20	13
Romania	16	11
Slovakia	91	42
Slovenia	19	15
Spain	27	15
Sweden	4	3
UK	2	1

Online services connected to MT@EC

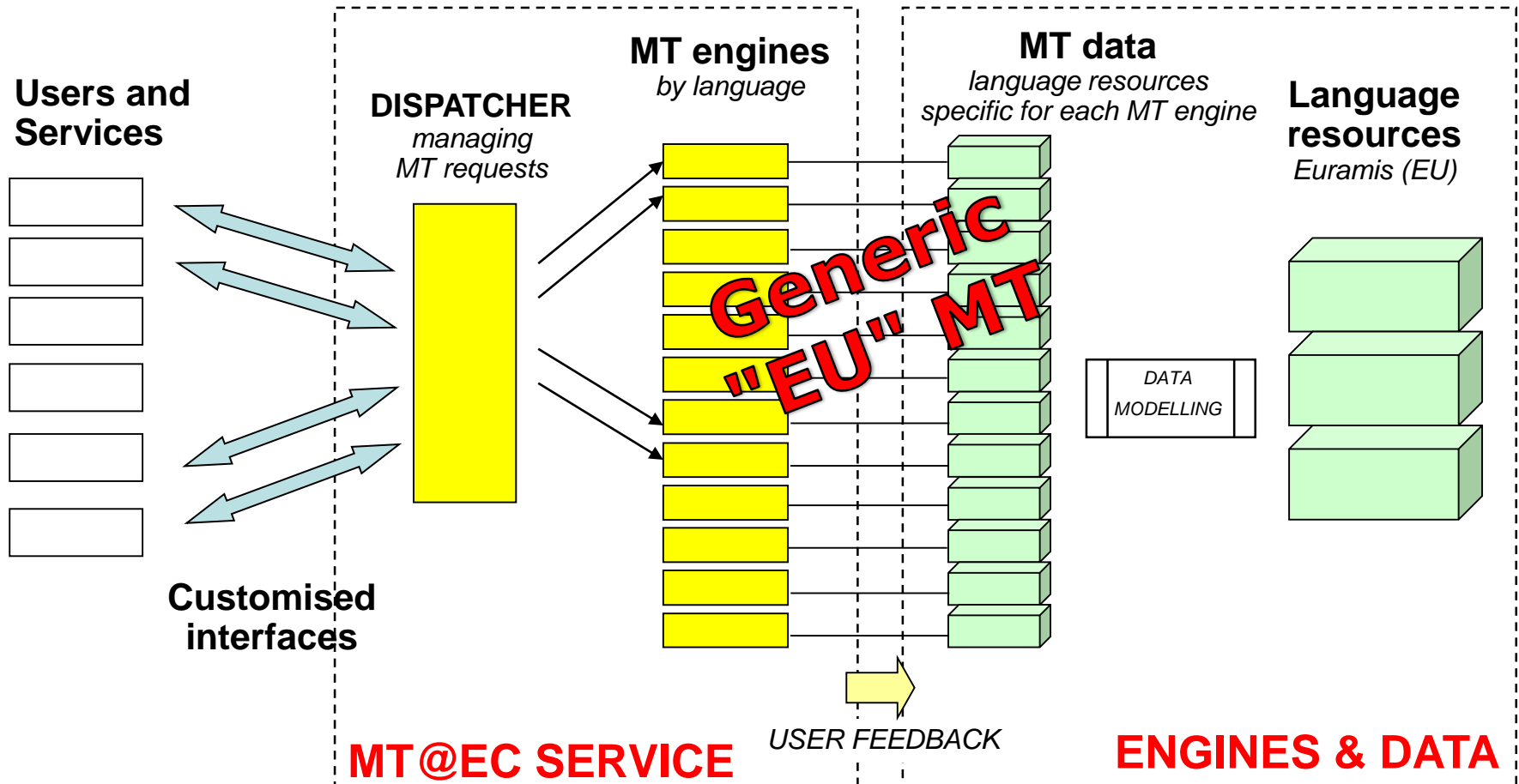
Service	Description/URL/use
IMI	Internal Market Information System – an intranet for more than 7000 public administrations in the EU Member States http://ec.europa.eu/internal_market/imi-net/index_en.html → free-text message exchange
SOLVIT	An on-line problem solving network concerning misapplication of Internal Market law by public authorities. http://ec.europa.eu/solvit/ → free-text message exchange
nLex	A common gateway to National Law http://eur-lex.europa.eu/n-lex/ → translate results of queries
CircaBC	Communication and Information Resource Centre for Administrations, Businesses and Citizens (collaborative spaces) https://circabc.europa.eu/ → Translate documents

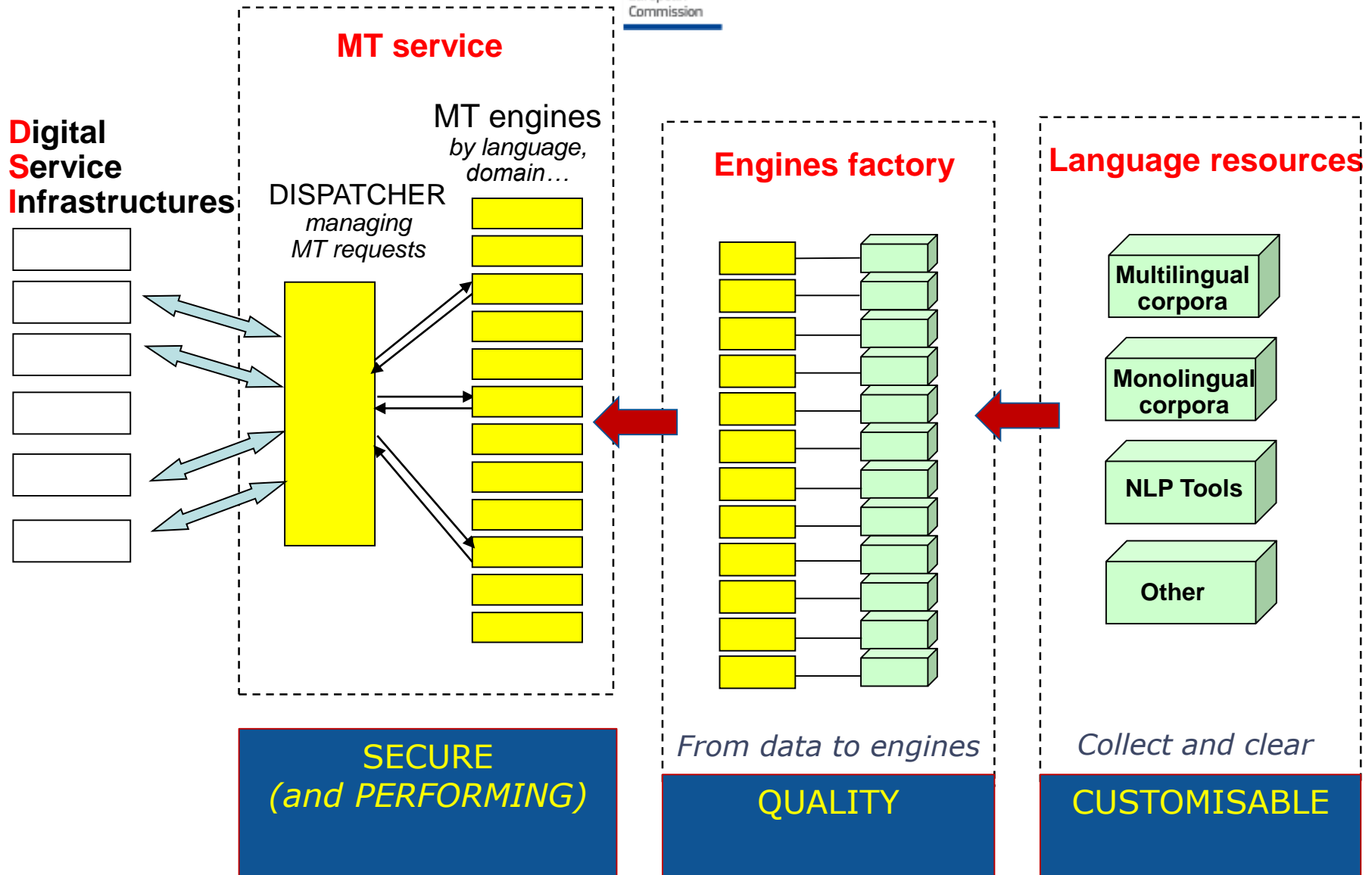
What next?

from MT@EC... to the CEF automated translation platform

CEF.AT will:

- build on the **existing** MT@EC service - but not be limited to it
- put emphasis on **secure, quality, customisable** MT for pan-European online services - but not be limited to them
- be a **multilingualism enabler** – not only MT





For CEF.AT we will be looking for:

- *even better **language technologies***
both open source and commercial products
- *more quality **language resources***
both open and licenced data
- *even better **ways to enable multilingualism***
both from research and market

*(and appropriate infrastructure to **scale up**)*

DGT contributes to this effort...

... as **solution provider** for the CEF.AT platform

... publishing **DGT-TM** a public translation memory
with almost **1.5 billion words** in **24 languages**

→ on the EU Open Data Portal

<http://open-data.europa.eu/data/dataset/dgt-translation-memory>

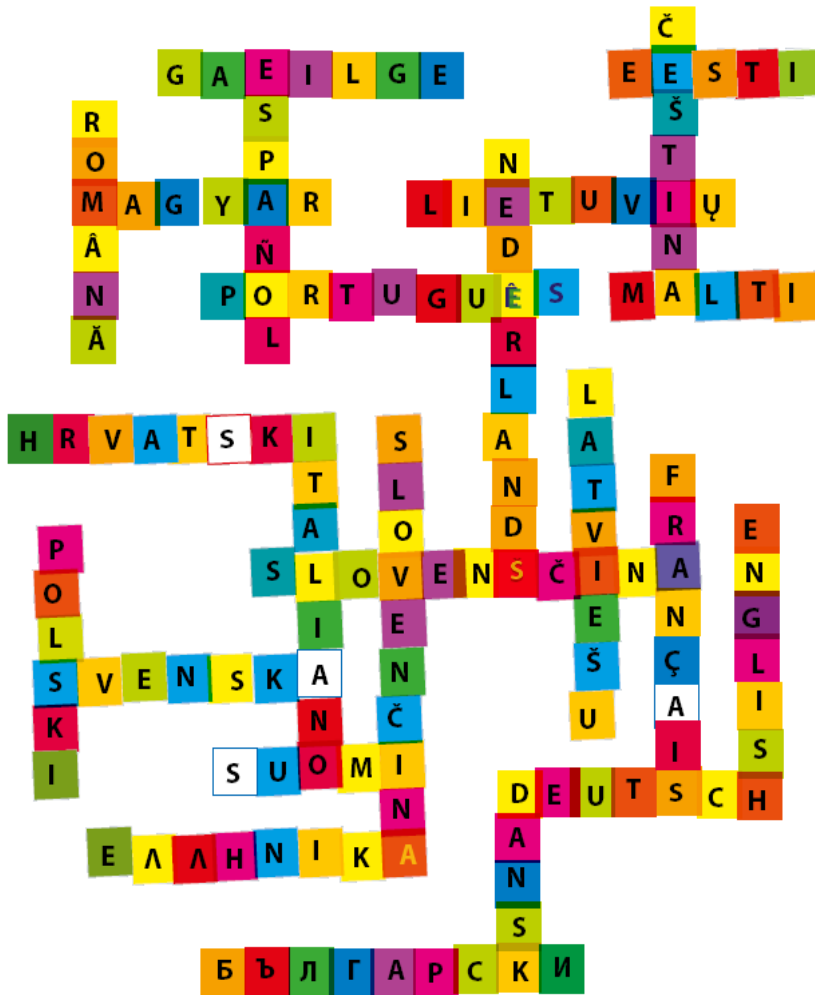
→ on the EC Joint Research Centre Language Technology web site

<http://open-data.europa.eu/data/dataset/dgt-translation-memory>

... supporting networking of translation stakeholders across the
EU through the **Translating Europe** initiative, aimed to create
a community of interest on translation

http://ec.europa.eu/dgs/translation/programmes/translating_europe/index_en.htm

The time for making Digital Europe multilingual
is **NOW**.



DGT events *for your agenda*

Annual Translating Europe Forum
Brussels, 29-30 October 2015
Focus: Youth

Annual MT@Work Conference
Brussels, 11 December 2015
Focus: MT in translation practice

[spyridon.pilos\(a\)ec.europa.eu](mailto:spyridon.pilos(a)ec.europa.eu)
[dgt-mt\(a\)ec.europa.eu](mailto:dgt-mt(a)ec.europa.eu)