



Language Resource Board

Virtual meeting, 29 April 2020

June Lowery-Kingston & Philippe Gelin

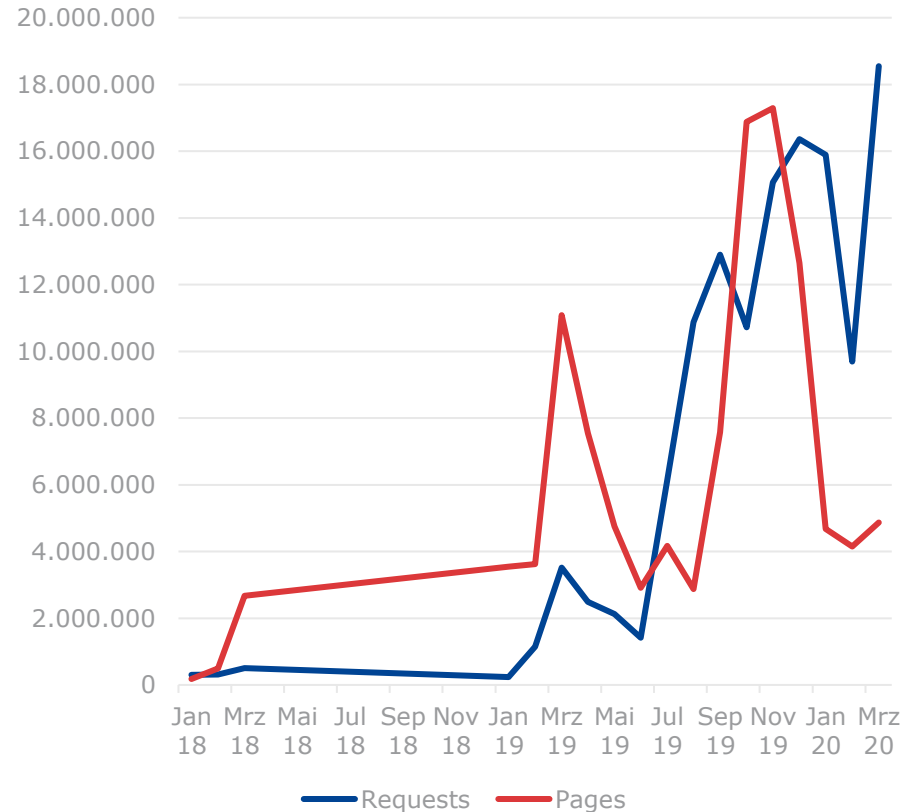
Head of Unit & Head of Sector Accessibility, Multilingualism & Safer Internet

Directorate-General for Communications Networks, Content and Technology
European Commission

eTranslation – Statistics

- Neural Machine Translation in the 24 official languages of the EU (24<->24) *plus Icelandic, Norwegian and Russian*
- Domain specific engines:
 - **Public health**, *culture, banking, IP Law, Case law, ... (not yet all 24<->24)*
- Used by more than 64 digital services
- Translation volume:
 - 27 Mio pages in 2018
 - 95 Mio pages in 2019
- Now available to SMEs across Europe

eTranslation usage (per month)



SME opening to eTranslation

- *"This tool is very timely for us because I have to prepare a skype meeting for next Wednesday with potential customers in Spain. I wanted to test the tool to automatically translate the presentation (from English to Spanish) and I am quite impressed with its performance and ease of use. So on my side it is a real added value."*
- *"I find it very useful to use the translation service, which has been very helpful in translating certain urgent documents into other languages."*

Registered users since last month



European Language Resource Coordination

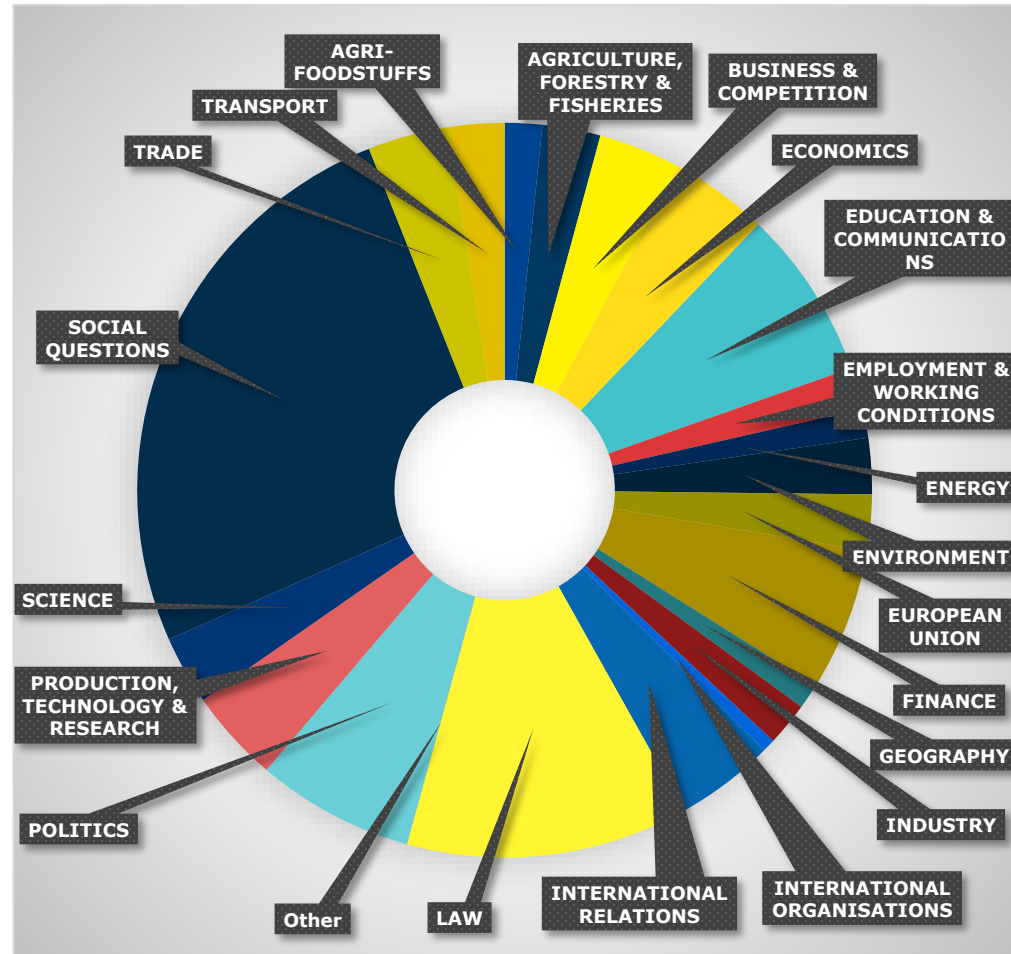
- 1367 language resources
- 3 billion words

COVID 19

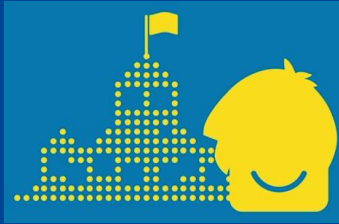
As of yesterday (28/4/20)

- 73 databases released

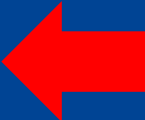
We need you!



CEF Telecom 2020 call for proposals



- Overall Budget: €4 million
- Indicative grants of €200,000 to €800,000 (5-20 projects)
- Co-funding: 75% of eligible costs
- Indicative duration: 2 years
- Deadline for submission: **Thursday 25 June 2020 (17:00.00 Brussels time)**



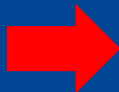
Three Priorities

1. Collect language resources
2. Collect language tools
3. Integrate and/or complement eTranslation into
 - CEF DSIs,
 - pan-European digital public services or
 - **Public administrations**
 - **SMEs**



Freely and openly available

Sustainable



Conditions for participation

- Consortia must be composed of
 - at least **two public or private entities**,
 - based in **two or more Member States** and/or EEA countries participating in CEF Telecom program (EU, Norway, Iceland)
- Proposal should address at least **two languages**
- Consortia members need prior Member State/EEA country agreement



No research

Examples of projects

- **Multilingual chat bot system**

A dialogue system that would cover two or more administrations in different member states that would provide useful information for cross-border workers, for example.

- **Cross-lingual automatic classification**

A classification system that would enable commerce SMEs commerce websites in Europe to automatically assign product categories with multilingual descriptions.

- **Speech synthesis** for languages for two or more low resources EU languages

Although speech synthesis technology has wide coverage, we still have EU languages with limited availability of systems. Freely available speech recognition would allow, for example, SMEs to have their webpages more accessible.

- **Speech transcription** for public administration and SMEs

2016	2017	2018	2019
eTranslation TermBank	Multilingual Resources for CEF.AT in the legal domain	Unleashing European Patent Translations (EuroPat)	Curated Multilingual Language Resources for CEF AT
ParaCrawl	National European Central Translation Memory Data (NEC TM Data)	CEF Data Marketplace	Federated eTranslation Termbank Network
ELRI	Broader Web-Scale Provision of Parallel Corpora for European Languages	Paracrawl III	Unsupervised MT for Low-resourced language pairs (MT4ALL)
Cross Border eProcurement	APE-QUEST - Automated Post-Editing & Quality Estimation (for EESSI and ODR)	PRINCIPLE (parallel data in HR, IS, IE and NO for eProcurement and eJustice)	Multilingual Anonymisation toolkit for Public Administrations (MA4PA)
EU Presidency	MICE (creating middleware providing MT-related services on top of the eTranslation output)	OCCAM (integration of image classification, TMs, OCR, and MT to support the Automated Translation of scanned documents)	User-focused Marian
IADAATPA		NTEU (Neural Translation for the European Union)	CEFAT4Cities
		EU Presidency II	

Acronym	EU Contribution (in EUR)	Project Start Date	Project End Date
Curated Multilingual Language Resources for CEF AT	720.000	01-06-2020	31-05-2022
Federated eTranslation Termbank Network	807.959	01-06-2020	01-06-2022
Unsupervised MT for Low-resourced language pairs (MT4ALL)	469.098	01-01-2020	31-12-2021
Multilingual Anonymisation toolkit for Public Administrations	912.620	01-01-2020	31-12-2021
User-focused Marian	608.879	01-03-2020	28-02-2022
CEFAT4Cities	785.246	01-04-2020	31-03-2022

More information on generic services projects: <https://ec.europa.eu/inea/en/connecting-europe-facility/cef-telecom/projects-by-dsi/etranslation?page=1>

Links & contacts:



CEF Call website:

<https://ec.europa.eu/inea/en/connecting-europe-facility/cef-telecom/apply-funding/2020-cef-telecom-calls-proposals>

eTranslation website:

<http://ec.europa.eu/cefdigital/etranslation>

eTranslation tool:

<https://webgate.ec.europa.eu/etranslation>

CNECT-LANGTECH@ec.europa.eu

June Lowery-Kingston HoU
Philippe Gelin HoS
Ainars Freimanis

In a nutshell

- Help us spread the word ...
...in all official languages

Identify,

Share,

Benefit!

***Now available to public administrations
across Europe***

Thank you!

Contact us



CNECT-LANGTECH@ec.europa.eu

© European Union, 2017. All rights reserved. Certain parts are licensed under conditions to the EU.
Reproduction is authorized provided the source is acknowledged.