



# 2018-2 CEF Telecom Call: CEF Automated Translation

Aleksandra Wesolowska  
Programme Officer – EU policies  
European Commission - DG CONNECT  
Unit G3 – Learning, Multilingualism & Accessibility

# What is CEF Automated Translation?



- CEF Automated Translation is a building block of the Connecting Europe Facility (CEF)
- The main purpose of CEF Automated Translation is to **make all Digital Service Infrastructure (DSIs) multilingual**
- CEF Automated Translation currently offers a **machine translation service (eTranslation)** but it is not limited to it (i.e. more MT technologies plus other NLP services can be part of the building block based on user needs)

Vision for CEF Automated Translation: **multilingualism enabler for public online services...**

## CEF Automated Translation building block today

**eTranslation service** [https://ec.europa.eu/info/resources-partners/machine-translation-public-administrations-etranlation\\_en](https://ec.europa.eu/info/resources-partners/machine-translation-public-administrations-etranlation_en)

- a central **machine translation service** run by Directorate-General for Translation (DGT)
- machine translation between 24 official EU languages, Icelandic and Norwegian (Bokmål)
- a machine translation service for public administrations and public online services in various domains/sectors
- available via a **web-user interface** and an **API**
- intended for the following types of users:
  - **online service providers** in European and national public administrations
  - **translators** in EU institutions and in national public administrations
  - **non-translating staff** in EU institutions and in national public administrations

**ELRC-SHARE repository service** <https://elrc-share.eu/>:

- repository of language resources needed to develop multilingual systems and services
  - language resource sharing infrastructure
  - part of the broad Europe-wide effort on **European Language Resource Coordination**
- [www.lr-coordination.eu](http://www.lr-coordination.eu)

*eTranslation v1.0 launched on 15 November 2017*

*eTranslation web page:*

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

*Web service (API) documentation available on request*

- *a full set (EN ↔ ALL) of speed-optimised engines*
- *neural MT engines for some languages:*
  - EN ↔ DE, HU, ET and FI,
  - EN → BG, CS, GA, LT, LV, PL, SK, SL
- *the median time for text snippets (1-5000 characters) is 1.5" seconds; for documents 2 minutes*
- *engines trained on EU-domain data*
- *domain-specific engines will be provided when sufficient data is available to DGT*





## **Ongoing activities supporting networking and language resource collection**

### **European Language Resource Coordination**

[www.lr-coordination.eu](http://www.lr-coordination.eu)

- networking and deepening collaboration with Member States
- conferences and workshops
- technical helpdesk and cleaning of language resources
- legal helpdesk and legal clearance of language resources
- onsite assistance for contributors of language resources
- ELRC-SHARE Repository: <https://elrc-share.eu/>

# CEF-TC-2018-2 Automated Translation call for proposals

- Budget: 5 M€
- Co-funding rate: 75% of eligible costs
- Pre-financing: 50% of the grant amount
- Indicative duration of the actions: 24 months

Two action lines:

- *Language resource projects: projects for facilitating language resource identification, processing and collection in the Member States*
- *Integration projects: to make online services multilingual with existing technology*

# CEF-TC-2018-2 Automated Translation Objectives

- (1) improvement of quality and coverage of automated translation in the public sector use cases relevant for CEF
- (2) boost the take-up of automated translation services among the CEF DSIs and other national online services relevant to those DSIs
- **No research activities**
  - **No development of tools or language resources**

## Objectives of "language resource" projects

- facilitate language resources identification, collection and processing in/by/with Member States
- addressing bottlenecks in language resource provision
- language resources: large bilingual or multilingual corpora, large monolingual corpora, domain-specific parallel corpora and terminologies...
- ... for under-resourced languages...
- ...in the fields of consumer rights, culture, legal domain, social security, health, public procurement, etc.
- building on the earlier work, in particular ELRC, but not duplicating ELRC effort nor ongoing CEF "LR projects"
- prioritization of language resources in the call text



## Objectives of "integration" projects

- work on a specific CEF DSI and/or national online services and portals related to CEF DSIs and integrate automated translation technologies
- targeted services:
  - **CEF DSIs**
  - **national online services and portals related to CEF services**
- use eTranslation (eTranslation must be integrated into a CEF-DSI-relevant service)
- application of other language technologies is possible and prioritized
- foster the multilingual best practice and architectures in public services

# CEF-TC-2018-2 Automated Translation: Admissibility and Eligibility

- choice of languages based on digital service's needs – at least two different EU/CEF languages
- eDelivery deployment may be part of the proposal
- consortium composition: consortium consisting of at least two entities from two or more Member States

## Objectives – differences to last year's call

- Applicants should address only **one** of the objectives indicated in section 2.1
- Prioritization for LR projects on (1) "**under-resourced**" languages in general and on (2) language resources in the CEF DSI domains
- **Obligation** for LR projects **to deliver resources to ELRC-SHARE** spelled out in call text
- Prioritization for integration projects which apply **language technologies other than machine translation**
- Other small differences in the wording of the call text

# award criteria – "language resources projects"

- **Relevance**
  - relevance of language resources for CEF Automated Translation (e.g. CEF DSI domains, under-resourced languages, CEF DSI domains, meaningful sizes)
  - alignment with European and national language policies, with clear differentiation from the ELRC activities
- **Quality & efficiency**
  - clarity on the ownership and licensing status of LRs
  - clarity on how privacy and/or legal issues have been/will be resolved
  - support from data-holders and national decision-makers (if promised resources: established contacts, letters of support, consortium members...?)
  - quality and effectiveness of the national networks and outreach
- **Impact & sustainability**
  - licensing status/"openness" of the language resources provided
  - the potential of the collected resources for various Language Technologies
  - contribution to creating a permanent & sustainable pipeline of language resources flowing to CEF Automated Translation platform

# award criteria "integration projects"

- **Relevance**
  - relevance/link of the addressed online service to CEF DSI(s)
  - alignment with European and national language policies and/or digital priorities
    - **Quality & efficiency**
      - maturity and readiness of the addressed digital service in terms of clarity on user requirements for multilinguality
      - clarity on the architectural adaptations needed to support multilinguality
      - clarity on how the integration with eTranslation will be established
      - availability and maturity of the other/specific LT tools/services to be integrated
      - appropriate involvement of the technology providers and owners of the digital service
        - **Impact & sustainability**
          - engagement and commitment of the owners of the CEF DSI/the digital service
          - Potential for reuse of the project results for other online services

## More information on the calls...



**inea-cef-telecom-calls@ec.europa.eu**  
**inea@ec.europa.eu**



**<https://ec.europa.eu/inea/en/connecting-europe-facility/cef-telecom/apply-funding/2018-cef-telecom-calls-proposals>**



**@inea\_eu**



**INEA**

**Thank you!**