



CEF Automated Translation DSI – setting the scene

Aleksandra Wesolowska
Programme Officer - EU policies
Data Value Chain Unit, DG CONNECT

CEF - Towards a Connected Multilingual Europe, 29/04/2015

What is CEF.AT?

- Secure and adaptable Automated Translation platform that plugs in to any DSI and makes it multilingual
- Builds on MT@EC machine translation system already in operation, serving EC institutions, some online services & portals and available to MS administrations
- CEF.AT is a broader concept: eventually platform that offer tools and services for multilingual enablement of public services

Features and enabling technologies of CEF.AT

- **START:** Adaptable machine translation and relevant Language Resources are central
- **FUTURE** likely enabling technologies: CAT, CMS, terminology management, semantic interoperability...
- **OTHER** key elements: interfaces to various DSIs, human intervention (helpdesk for DSI, quality control & validation, post-editing of MT output, on-demand response...)

CEF.AT in a nutshell

- Automated Translation (AT) is a "building block"
- AT will serve the other Digital Services Infrastructures in CEF
- AT = whatever it takes to make DSIs actually multilingual

CEF.AT – actions in work programme 2014

Requirements analysis

- analysing and documenting the multilingual requirements of CEF DSIs
- Designing the IT & service infrastructure for CEF.AT (architecture, data processing and storage environment)

Deployment of small-scale systems

- adapting, interfacing and deploying MT@EC system to three CEF DSIs

Language resources coordination

- training workshops for Member States and other stakeholders, to establish a functional network of providers of language resources for CEF.AT
- technical secretariat and helpdesk for the language resource coordination and involvement of MS's and other stakeholders

CEF.AT – actions in work programme 2015

Core CEF.AT platform

- *IT and software resources* for scaling up CEF.AT to realistic scale of use, serving 5-6 DSIs, based on requirement analysis and small-scale deployment
- *language resource repository* to store and process language resources obtained from MSs and other contributors
- *adaptation factory* that automates the adaptation of CEF AT to new text types and use situations

Language resources for automated translation

- extend and deepen the *coordination* set up in WP 2014, move from training and networking to systematic *collection* of language resources
- Provide an *IPR support* and *clearance* helpdesk

Why Language Resource Coordination?

- This is the first operational event of the CEF.AT language resource coordination efforts
- CEF.AT needs parallel text corpora, terminologies, vocabularies, taxonomies – the "raw material" for automated translation technology
- Focus and concentrate on resources relevant for CEF
- CEF.AT is joint effort and shared responsibility



Thank you!

Email: aleksandra.wesolowska@ec.europa.eu