THE CEF AUTOMATED TRANSLATION PLATFORM

FRANÇOIS THUNUS, DGT
11th June 2021

A BIT OF HISTORY

MT@EC
- Legalese
- Statistical MT

eTranslation
- Neural MT
- Domain adaptation
- Integration with online public services

CEF.AT
- More LT tools, not just translation
- Supporting actions for data collection – ELRC
- Supporting funded projects

WHO'S IT FOR?

ONE SYSTEM – TWO AUDIENCES

- CEF Digital Service Infrastructures
- Pan-European digital public services
- Public administrations in Member States, Iceland and Norway
- SMEs

CEF eTranslation

- Translators and staff of the EU Institutions
- Digital services of the EU Institutions

DGT eTranslation

WHAT CAN I DO WITH ETRANSLATION?

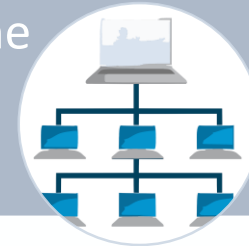# ONE SYSTEM – TWO USE SCENARIOS

- Automatically translate your texts for review by translators
- Rough translations to get the gist

Web interface for humans

- Integrate in workflows, websites, digital services, etc.

Machine-to-machine service (API)

Integrate eTranslation directly into your digital services

Request machine translations for text snippets or documents on demand

Qualifications translated

Download

Translations to and from any of the 24 EU official languages, Icelandic and Norwegian

Translate documents in various formats (doc, pdf, odt, xlsx, html etc.)

eTranslation now available for Russian, Mandarin, Japanese and Turkish!
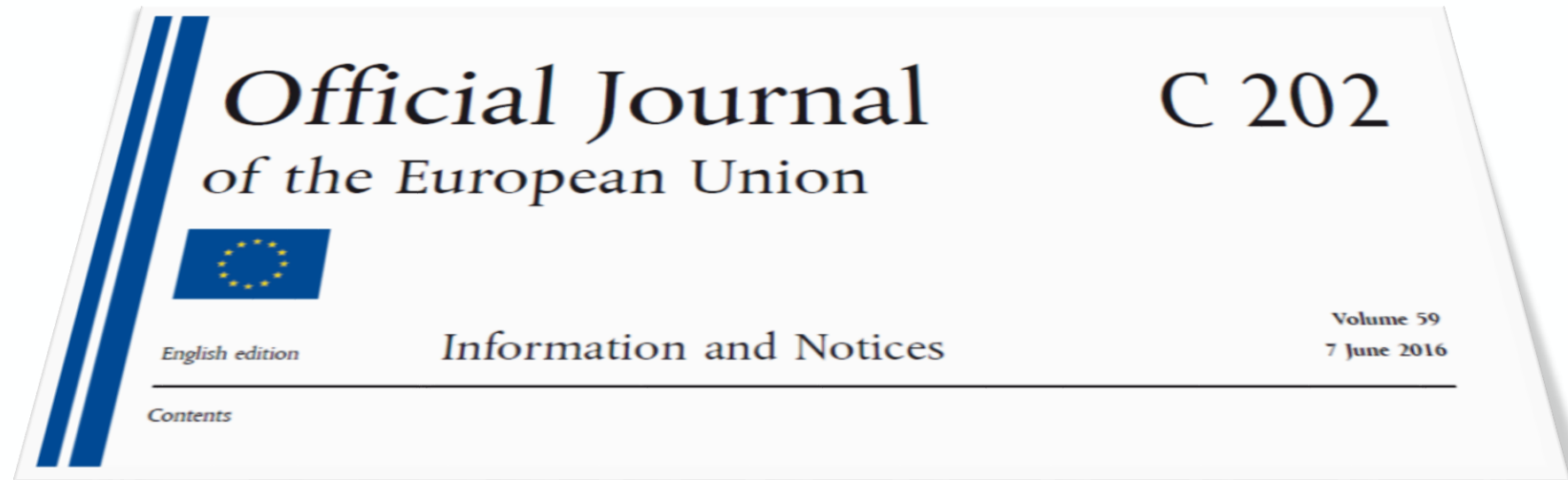
And more non-EU languages to come…

DOMAINS COVERAGE

- EU formal language
- General text
- Court of Justice Case Law
- Cultural
- Deutsche Bundesbank
- IP Case Law
- Ministère des Finances (France)
- Public Health
- Technical Regulation Information Systems
- Valtioneuvoston Kanslia

WHAT ABOUT THE TRANSLATION QUALITY?

WHAT WORKS BEST

- Texts related to EU policies

….because eTranslation is built using translated EU documents and it reuses them

| Translate text | Translate documents | My translation requests | | My settings | English | Logout |

the price of failure is too high.
energy is the life blood of our society

73 / 2500

falen is geen optie - de prijs is te hoog
energie is de levensader van ons samenleving

Energy 2020
A strategy for competitive, sustainable and secure energy

Energie 2020
Een strategie voor een concurrerende, duurzame en continu geleverde energie

INTRODUCTION

The price of failure is too high.

Energy is the life blood of our society. The well-being of our people, industry and economy depends on safe, secure, sustainable and affordable energy. At the same time, energy related emissions account for almost 80% of the EU's total greenhouse gas emissions. The energy challenge is thus one of the greatest tests which Europe has to face. It will take decades to steer our energy systems onto a more secure and sustainable path. Yet the decisions to set us on the right path are needed urgently as failing to achieve a well-functioning European energy market will only increase the costs for consumers and put Europe's competitiveness at risk.

INLEIDING

Falen is geen optie - de prijs is te hoog ...

Energie is de levensader van onze samenleving. Het welzijn van onze bevolking, industrie en economie hangt af van een veilige, continu geleverde, duurzame en betaalbare energie-voorziening. Tezelfdertijd zijn energiegerelateerde emissies goed voor bijna 80% van de totale uitstoot van broeikasgassen in de EU. Het energieprobleem vormt een van de grootste actuele uitdagingen voor Europa. Het zal decennia duren om onze energiesystemen een meer veilige en duurzame richting uit te sturen. Maar de beslissingen die ons op de goede weg moeten brengen, moeten vandaag al worden genomen. Als we er niet in slagen een goed functionerende Europese energiemarkt tot stand te brengen, zal dit de kosten voor de consument doen oplopen en het concurrentievermogen van Europa in het gedrang brengen.

WHAT WORKS LESS WELL

- Non-standard, new or creative text
- Single words or expressions
- Anything that depends on context

AN EXAMPLE: ETRANSLATION ON NON-STANDARD TEXT

EE formal language VS. General text engine

WHAT'S COMING?

- extending domain coverage (e.g. scientific texts; social media)

- extending language coverage (non-EU languages of social & economic importance, regional languages)

- additional language technologies
  - anonymization
  - basic CAT tool
  - (named-entity recognition)
  - (speech-to-text)

https://language-tools.ec.europa.eu/

eTranslation	Multilingual Tweet	Speech-to-Text	Named-entity Recognition
Interactive Terminology for Europe	European Language Resource Coordination (ELRC)	Catalogue of services	CEF Building Block Information

HOW TO REGISTER?
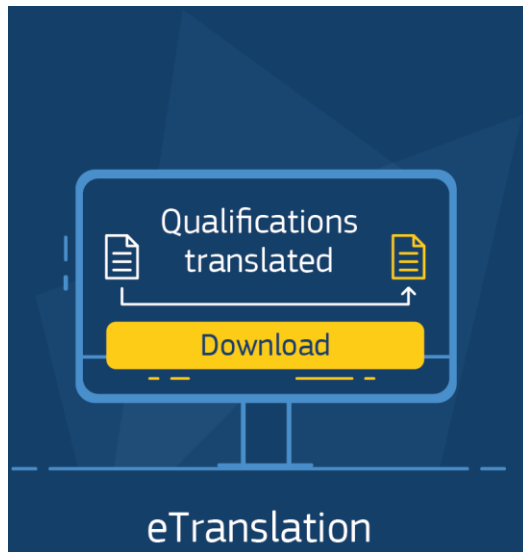
- Eligible individual users
  - Public Administrations
  - Universities
  - CEF-funded projects
  - SMEs (upon validation)
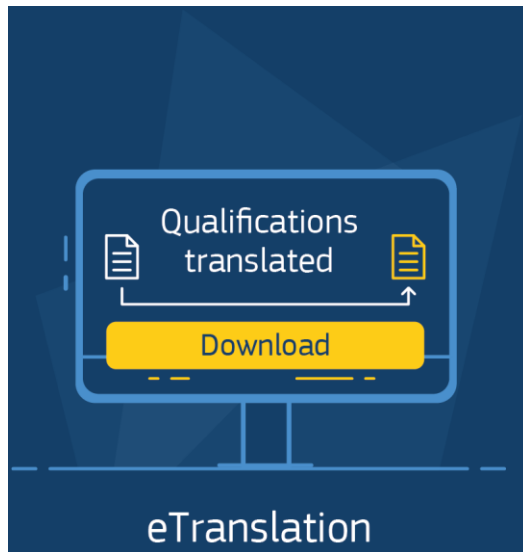
- Eligible individual users
  - Self-registration via https://webgate.ec.europa.eu/etranslation/public/welcome.html

- Web service (API) integration
  - Technical documentation: https://ec.europa.eu/cefdigital/wiki/display/CEFDIGITAL/How+to+submit+a+translation+request+via+the+CEF+eTranslation+webservice
  - Contact the eTranslation **Service Desk** for credentials (and other support): help@cefat-tools-services.eu

- Questions asked:

- 1) - Can CEF eTranslation also be used through API so that it can be integrated in other applications, e.g. a chatbot of the Dutch councils.

- 2) With regard to machine translation, to what extent can we rely on the correct translation and therefore correct output?

- 3) How is the output of the CEF Tools tested?

- 4) When is translation with AI desirable, when not and which disclaimer is appropriate?

- 5) Security

- 6) Mp3 format

GIVE IT A TRY!

HTTPS://LANGUAGE-TOOLS.EC.EUROPA.EU

DANK U VOOR UW AANDACHT!