



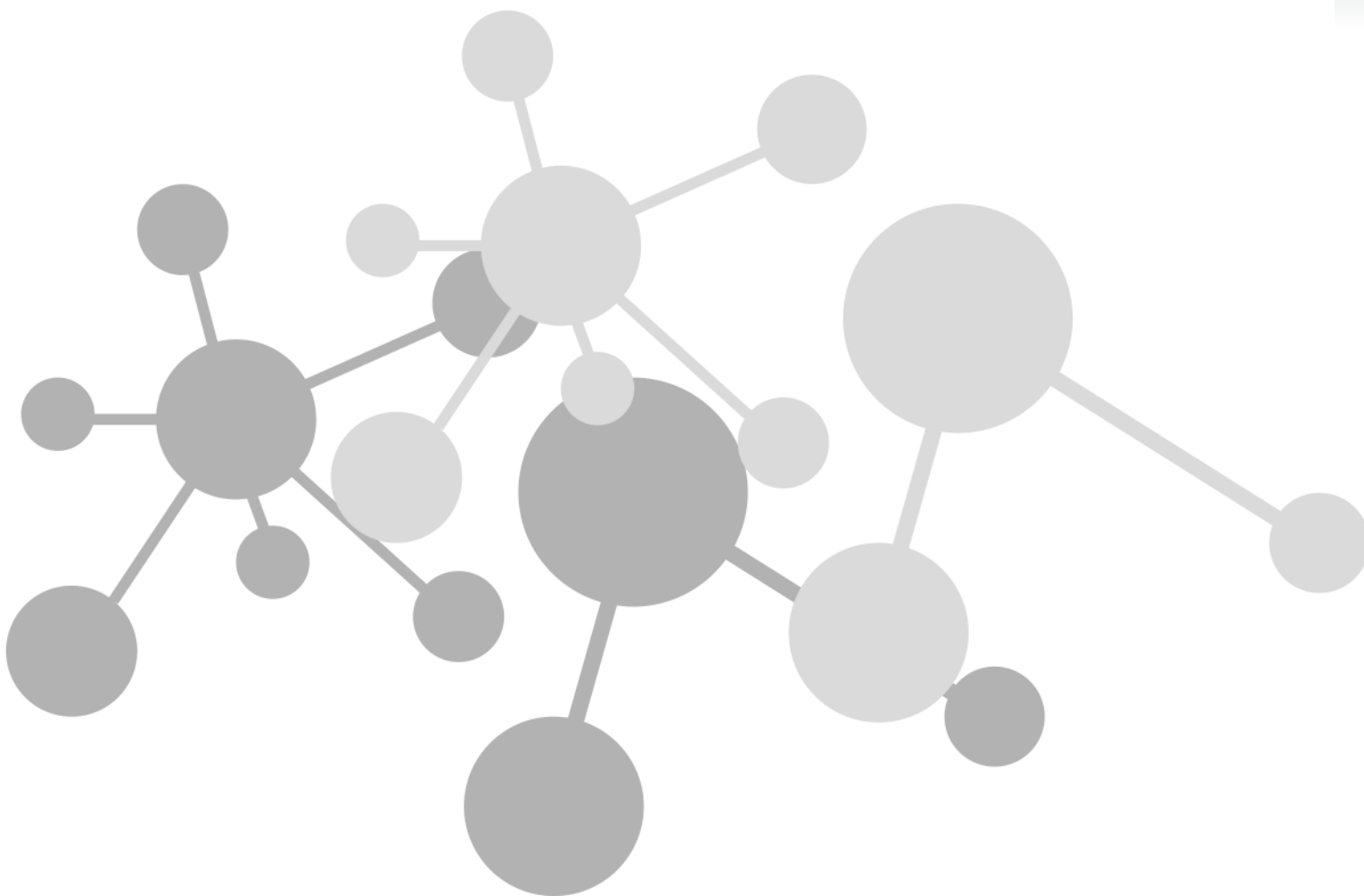
# Terminology for Europe: eTranslation TermBank

Tatjana Gornostaja

Tilde



Co-financed by the European Union  
Connecting Europe Facility



# eTranslation TermBank

Stimulates the collection and provision of **terminological resources** in sector-specific domains and languages of interest to sector-specific digital public services

## Sector-specific domains and digital public services in focus:

- Health to support CEF DSI **eHealth**
- Business legislation to support CEF DSI **e-Justice**
- Customer protection to support CEF DSI **ODR**

## Outcomes:

- 150 terminological resources for **official languages of the EU** plus **Norwegian and Icelandic**
- "Terminology for Europe" network for **sustainable cooperation**
- Pipeline of **terminological resources flow to CEF.AT**
- Directory of entities dealing with **terminology on national level**
- Directory of existing **terminological resources in** the three **sector-specific domains** for CEF.AT
- Methodology for **preparation and application of terminological resources** for CEF.AT

## Impact:

- Improve the **quality and coverage of CEF.AT**

## Partners:

**Tilde**, Latvia (coordinator)  
**University of Copenhagen**, Denmark  
**Árni Magnússon Institute for Icelandic Studies**, Iceland  
**Institute of the Estonian Language**, Estonia  
**Jožef Stefan Institute**, Slovenia  
**International Network for Terminology**, Austria  
**Institute for Language and Folklore**, Sweden  
**Swedish Centre for Terminology**, Sweden  
**Institute of the Lithuanian language**, Lithuania



Institut "Jožef Stefan", Ljubljana, Slovenija



Institute of the Estonian Language



Institute of the Lithuanian language

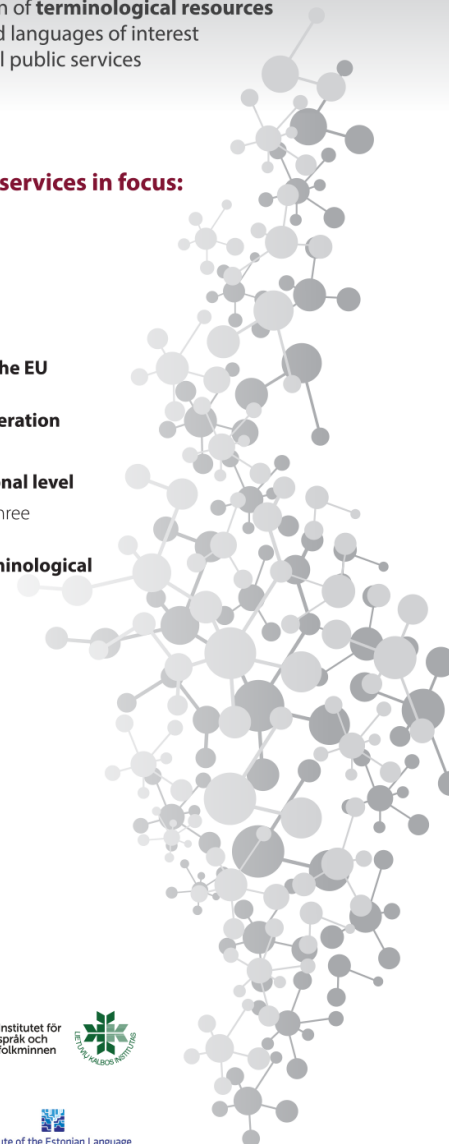


Institute of the Lithuanian language



Co-financed by the European Union  
Connecting Europe Facility

Project eTranslation TermBank is co-financed by the European Union's Connecting Europe Facility

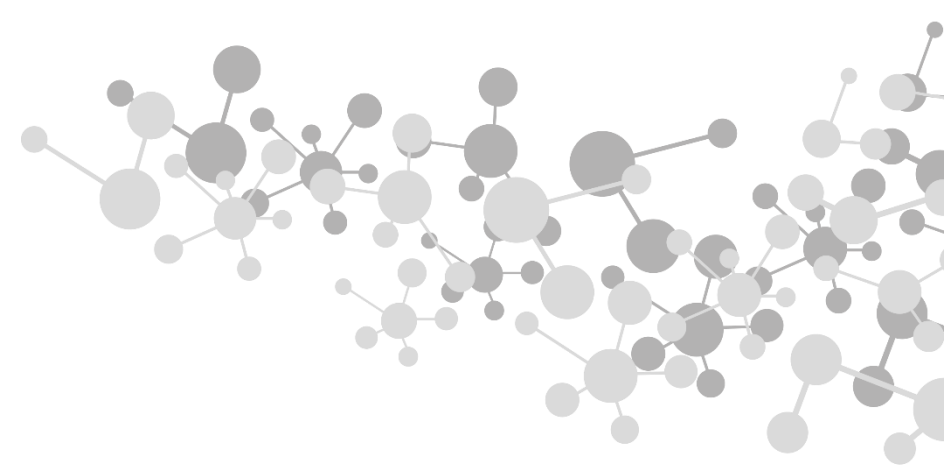




# Language Resource project: definition

...collect and provide **terminological resources of high quality, with clear licensing status, in specific domains appropriate for sector-specific Digital Service Infrastructures, and in languages of interest**

in order to improve the quality and coverage of automated translation offered by the CEF eTranslation service, maintain and advance its improvement over time, based on mature operational and technical solutions implemented during the project, and thus boost the take-up of existing and future CEF eTranslation services



# Why terminological resources for automated translation?



# Terminology for automated translation

Current state of the art machine translation is based on **AI technologies**

**neural machine translation (NMT)** uses a large artificial neural network  
**performs better in most cases** than previous paradigm of statistical machine translation (SMT)

while **SMT** does **better for out of domain** content, **NMT** likes things that are **domain-specific**

nevertheless, **terminology matters for both SMT and NMT**

consider **virus** in **medicine** and **computer science**

both systems need **terminology to disambiguate**



# eTranslation TermBank project: goal

eTranslation TermBank aims at providing **new terminological resources** to the language resource pipeline of CEF eTranslation for **three specific domains**

Health

Business legislation

Customer protection

to support the three sector-specific **CEF Digital Service Infrastructures**

eHealth

e-Justice

Online Dispute Resolution



# Terminological resource: definition

**New** terminological resources covering languages of interest – **all official languages** of the European Union plus **Norwegian and Icelandic**

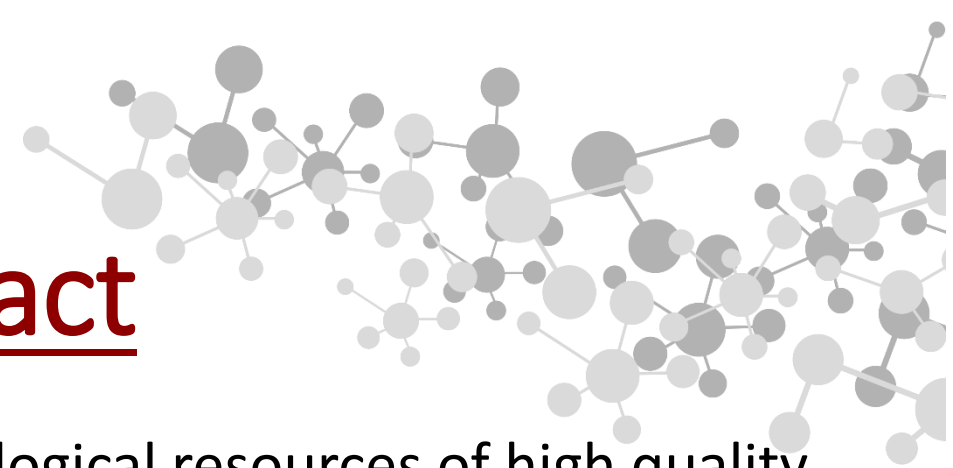
a **terminological resource** is a set of **monolingual, bilingual, or multilingual** terminological data in a machine-readable form,  
new terminological resources within the project are terminological resources, which are **not available via the ELRC-SHARE language resource repository** at the time and are **appropriate for sector-specific CEF DSIs** covered by the project



# eTranslation TermBank: outcomes

- A minimum of **150 terminological resources** for all CEF languages and in three specific domains for CEF eTranslation
- Stakeholder network “**Terminology for Europe**” for sustainable cooperation with CEF eTranslation
- **Pipeline of terminological resources flow** to CEF eTranslation
- **Case study and terminology requirement analysis** of trans-European and national CEF Digital Service Infrastructures
- The **directory of entities dealing with terminology** on European, regional, and national levels and **directory of existing terminological resources** in the three specific domains for CEF eTranslation
- The **methodology for preparation and application of terminological resources** for CEF eTranslation:
  - selection, harmonisation, and consolidation
  - methodology to integrate terminological resources into automated translation





# eTranslation Termbank: impact

...to **stimulate the collection and provision** of terminological resources of high quality and with clear licensing status in specific domains and languages of interest to sector-specific CEF DSIs

in order to **improve the quality and coverage** of automated translation offered by **CEF eTranslation**, maintain and advance its **improvement over time** based, on **mature operational and technical solutions** implemented during the proposed action, and boost the take-up of **existing and future CEF eTranslation services**

...to raise awareness about the importance of terminological resources to **improve the quality and coverage of automated translation offered by CEF eTranslation**

# eTranslation TermBank

Stimulates the collection and provision of **terminological resources**  
in sector-specific domains and languages of interest  
to sector-specific digital public services



Tilde, **Latvia** (coordinator)

University of Copenhagen, **Denmark**

Árni Magnússon Institute for Icelandic Studies, **Iceland**

Institute of the Estonian Language, **Estonia**

Jožef Stefan Institute, **Slovenia**

International Network for Terminology, **Austria**

Institute for Language and Folklore, **Sweden**

Swedish Centre for Terminology, **Sweden**

Institute of the Lithuanian language, **Lithuania**

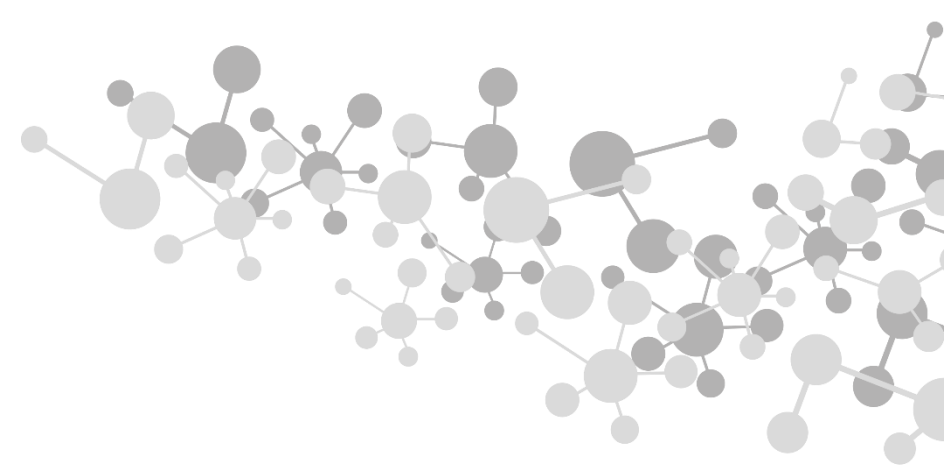


Co-financed by the European Union

Connecting Europe Facility



Co-financed by the European Union  
Connecting Europe Facility



# What is our mission?



# Raise awareness: visible vs. invisible terminology

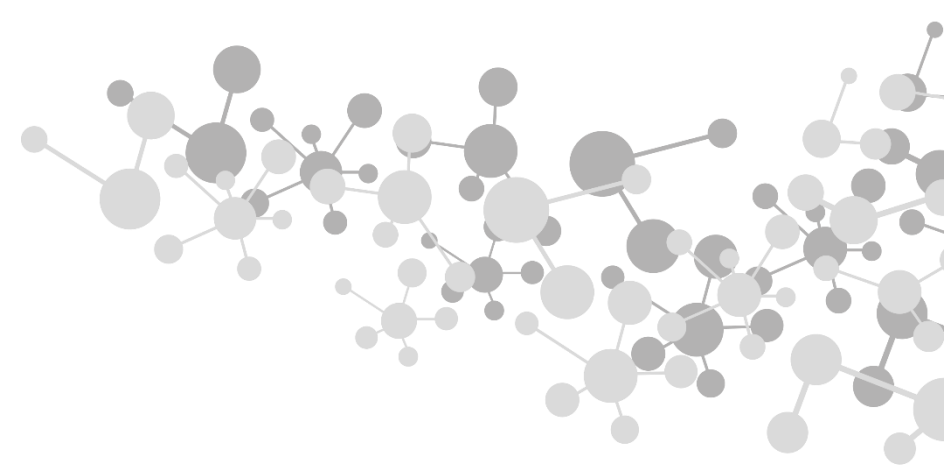
- Raise awareness about the importance of terminology for automated translation
- Introduce and support the definition of open terminological data – shall we?

"Open Data creates value and can have a positive impact in many different areas" ([Open Data Handbook](#))

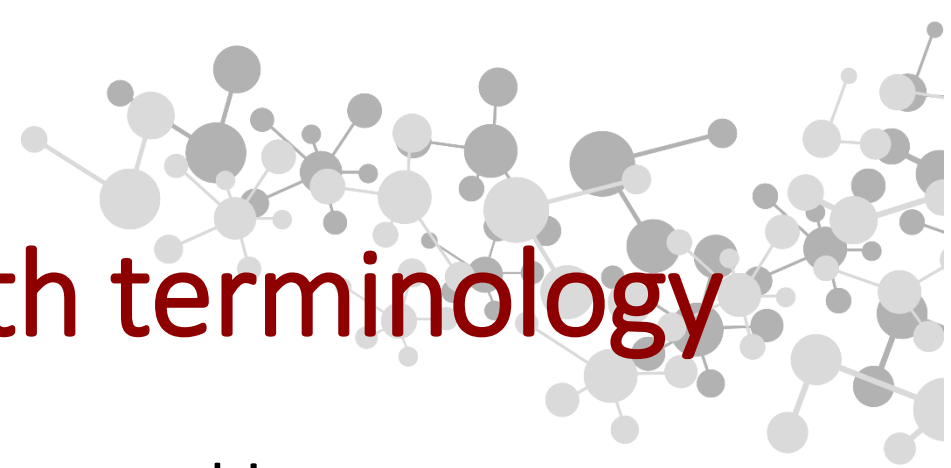
- Contribute to data interoperability

Ability of diverse systems and organisations to work together

Open and visible 😊



# Who can participate?



# All professionals working with terminology

Join “Terminology for Europe” network!

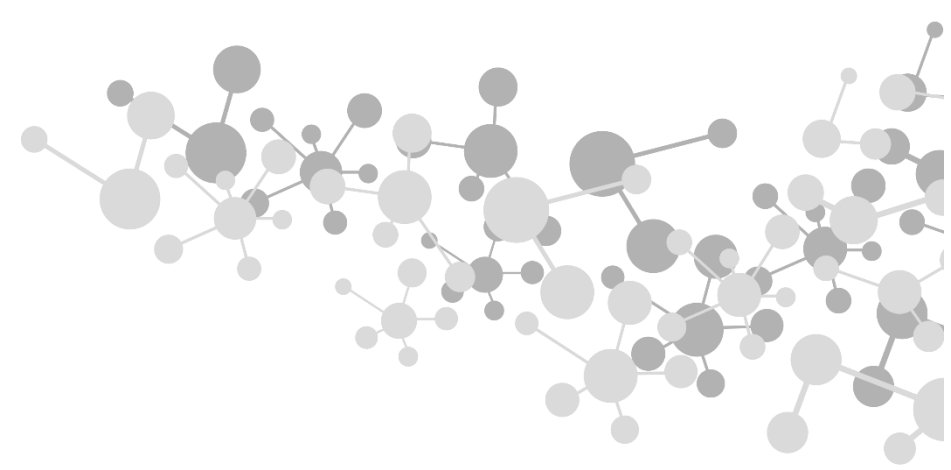
- Research institutions & students
- Public administrations and institutions
- Language service providers & freelancers
- Other businesses and initiatives

We need your involvement!



Co-financed by the European Union

Connecting Europe Facility



# Why to participate?



# The more we give, the more we get

Terminology matters and is vital for good translation quality!

- We share our terminological resources with colleagues
- We open our terminological resources for community
- We ensure our terminological resources are visible
- We contribute to good translation quality of automated translation
- We help to connect public administrations, businesses and citizens

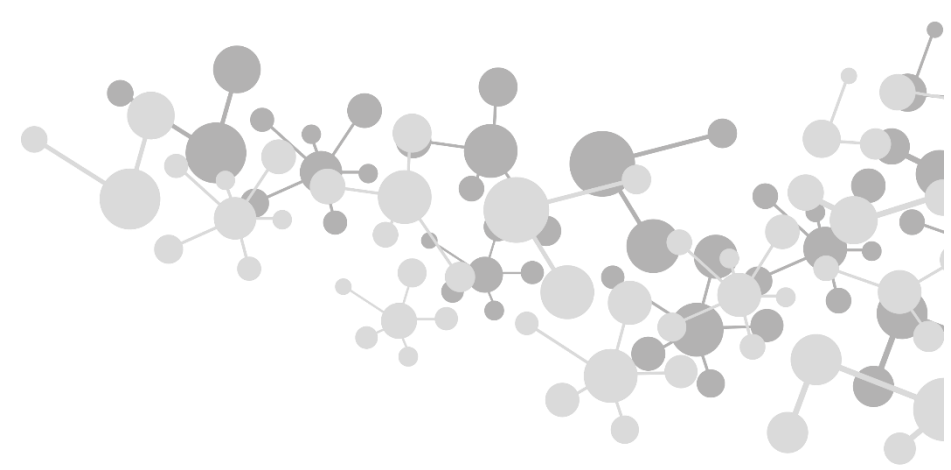
Do not underestimate the value of your terminology!





Co-financed by the European Union

Connecting Europe Facility



# How to participate?



# Share and open terminological resources

Monolingual, bilingual and multilingual terminological resources!

- For all EU languages, Icelandic and Norwegian
- In machine-readable form
- TBX is perfect, other formats are appreciated
- Various domains e.g. healthcare, legal & business, customer protection

Our power is in together 😊



Thank you 😊

[tatjana.gornostaja@tilde.com](mailto:tatjana.gornostaja@tilde.com)