



European Language Resource Coordination

ELRC Vilnius Workshop | 31.01.2019

Andrejs Vasiljevs | ELRC/Tilde



European
Commission



European Language
Resource Coordination
Connecting Europe Facility



How to provide smaller language communities with the same opportunities that the largest language communities enjoy?

62% of Europeans
don't speak English well enough
to hold a conversation

42% of Europeans **never**
purchase products or services in
a language other than their own

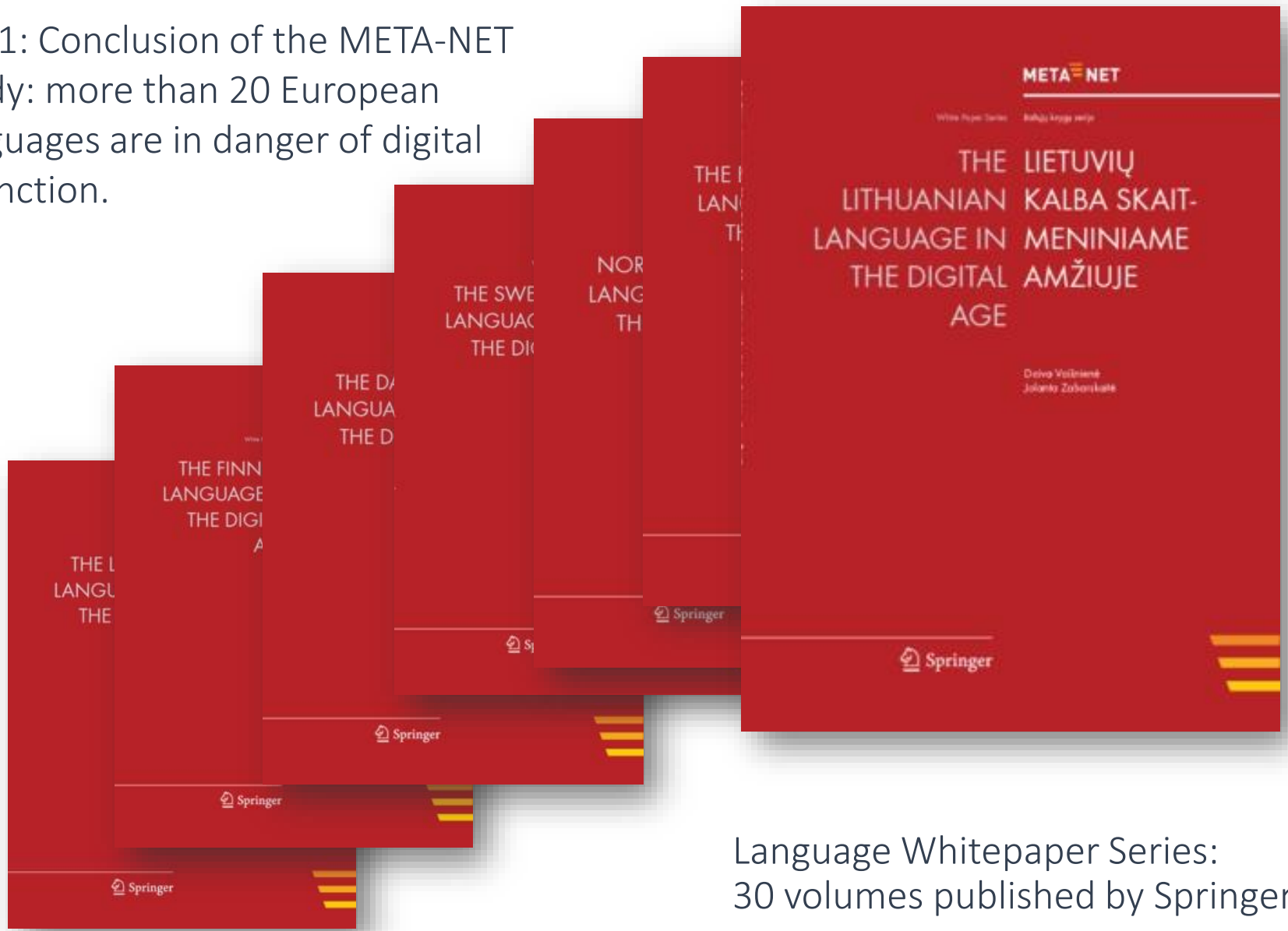


The Second Gutenberg effect

Languages that will not be freely used
in the digital technologies are going to
face **digital extinction**

2011: Conclusion of the META-NET study: more than 20 European languages are in danger of digital extinction.

META-NET



Language Whitepaper Series:
30 volumes published by Springer

Machine Translation support for European languages

Excellent support	Good support	Moderate support	Fragmentary support	Weak/no support
	<ul style="list-style-type: none">English	<ul style="list-style-type: none">FrenchSpanish	<ul style="list-style-type: none">CatalanDutchGermanHungarianItalianPolishRomanian	<ul style="list-style-type: none">BasqueBulgarianCroatianCzechDanishEstonianFinnishGalicianGreekIcelandicIrishLatvianLithuanianMalteseNorwegian (Bokmål, Nynorsk)PortugueseSerbianSlovakSloveneSwedishWelsh

eTranslation

Europe's automated translation platform

Highest translation quality for all CEF languages

Fully integrated into digital services and European portals

Used by **public administrations** in all CEF countries

Enables **secure exchange** of official documents

Benefits of using the BUILDING BLOCKS

REDUCE
COSTS

SHORTER
TIME TO
MARKET

CREATE A
DIGITAL
SINGLE
MARKET

Connecting Europe Facility with eTranslation

Exchanging information across languages in the
EU Member States, Iceland and Norway

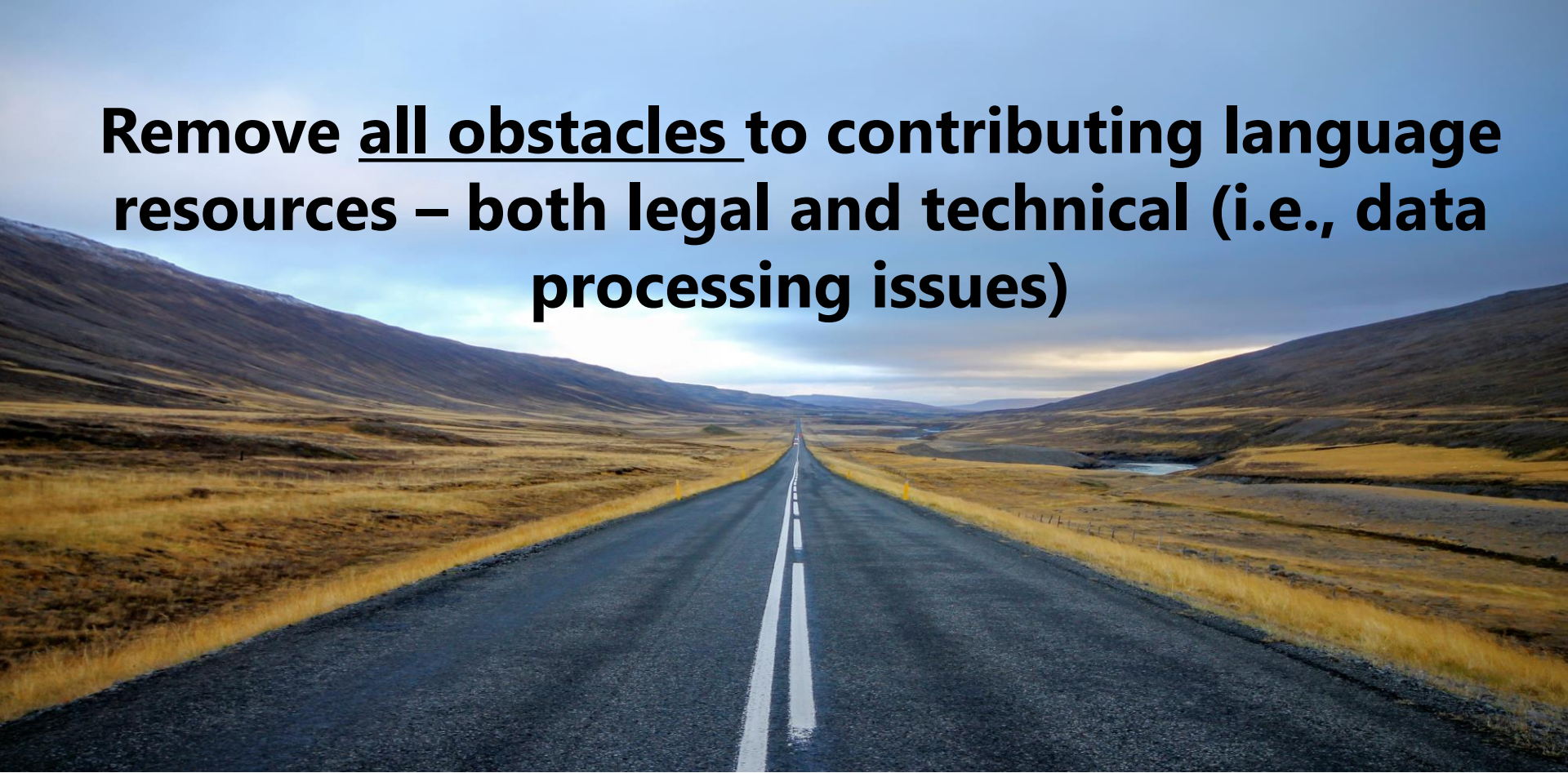


Why ELRC...?



Help data contributors to provide language resources for building better MT systems in CEF eTranslation, benefiting all of Europe





Remove all obstacles to contributing language resources – both legal and technical (i.e., data processing issues)

What does ELRC do?



Collect

Language resources

Identify

Needs of public services

Engage

With the public sector in the
identification of LR

Help

With any technical or legal issues

Act

Observatory for language resources
across Europe

Heads & organisations behind the ELRC



- ↳ 30 ELRC Technological National Anchor Points
- ↳ 30 ELRC Public Services National Anchor Points



30 ELRC Technological National Anchor Points

Bolette Sandford Pederse
*Centre for Language Technology, University
of Copenhagen*

30 ELRC Public Services National Anchor Points

Sabine Kirchmeier-Andersen
Danish Language Council

Peter Gylden Houmann
Agency for Digitisation

Count	Corpus	Lexical conceptual resource	Translation units
Danish	12	1	91 335
Estonian	6	1	128 401
Finnish	8		200 307
Icelandic	19	1	490 958
Latvian	5	1	66 550
Lithuanian	8	1	204 176
Norwegian	59	2	1 850 253
Swedish	15	4	103 057



German Centre for
Artificial Intelligence



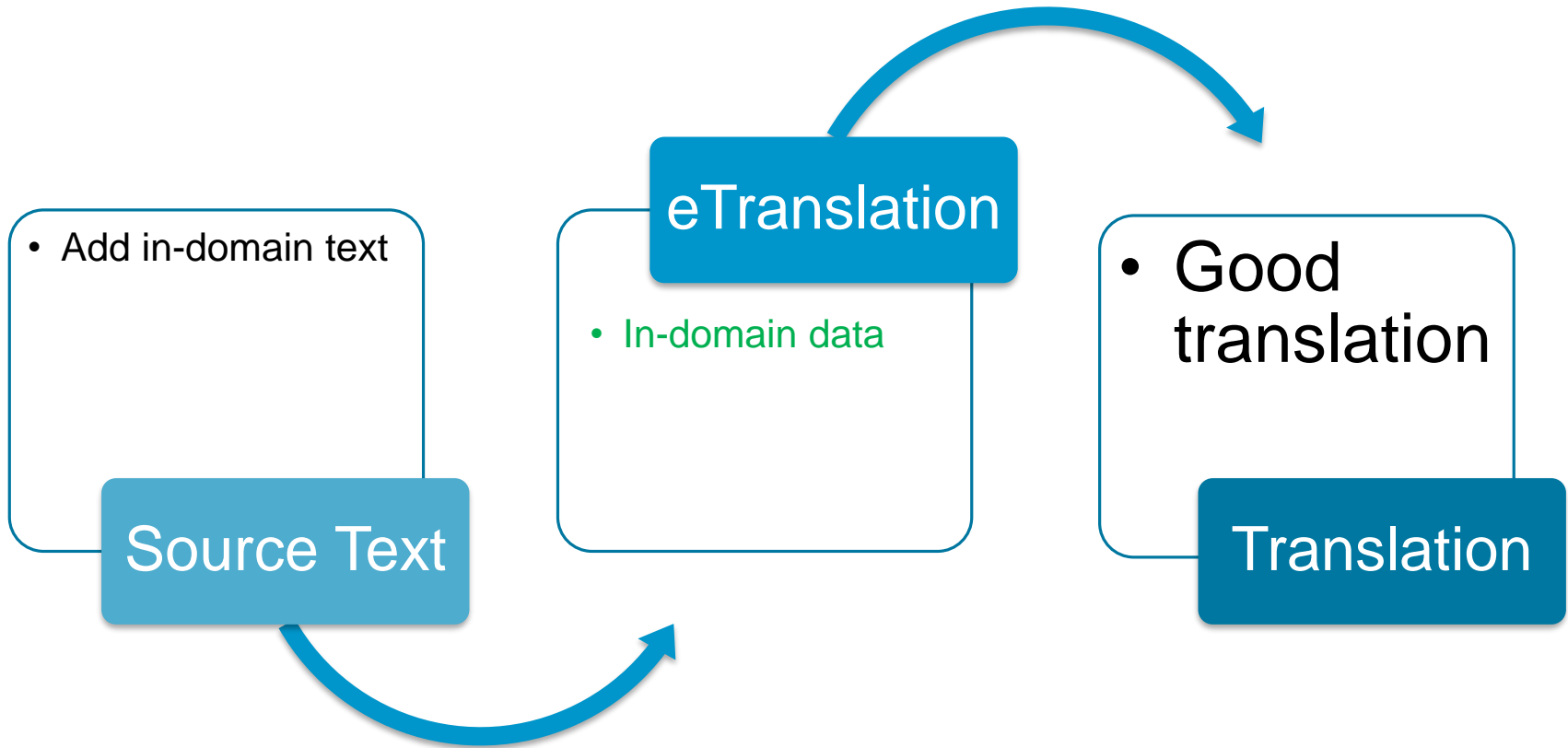
Institute for Language and
Speech Processing

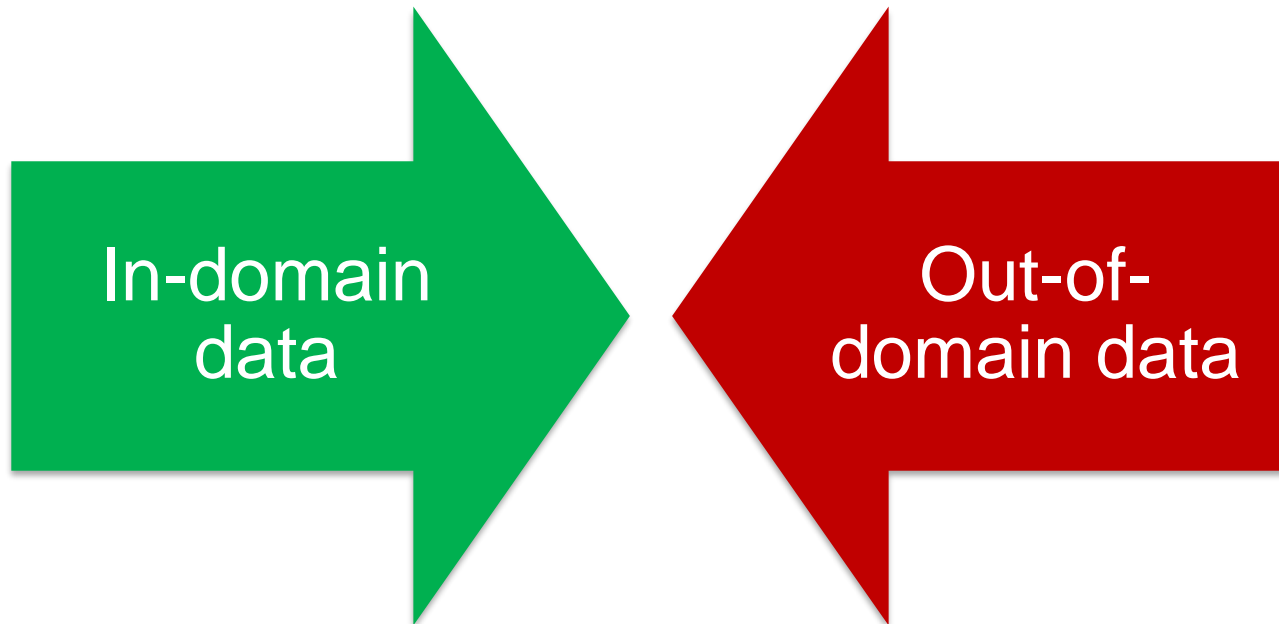


Tilde



European Language Distribution Agency

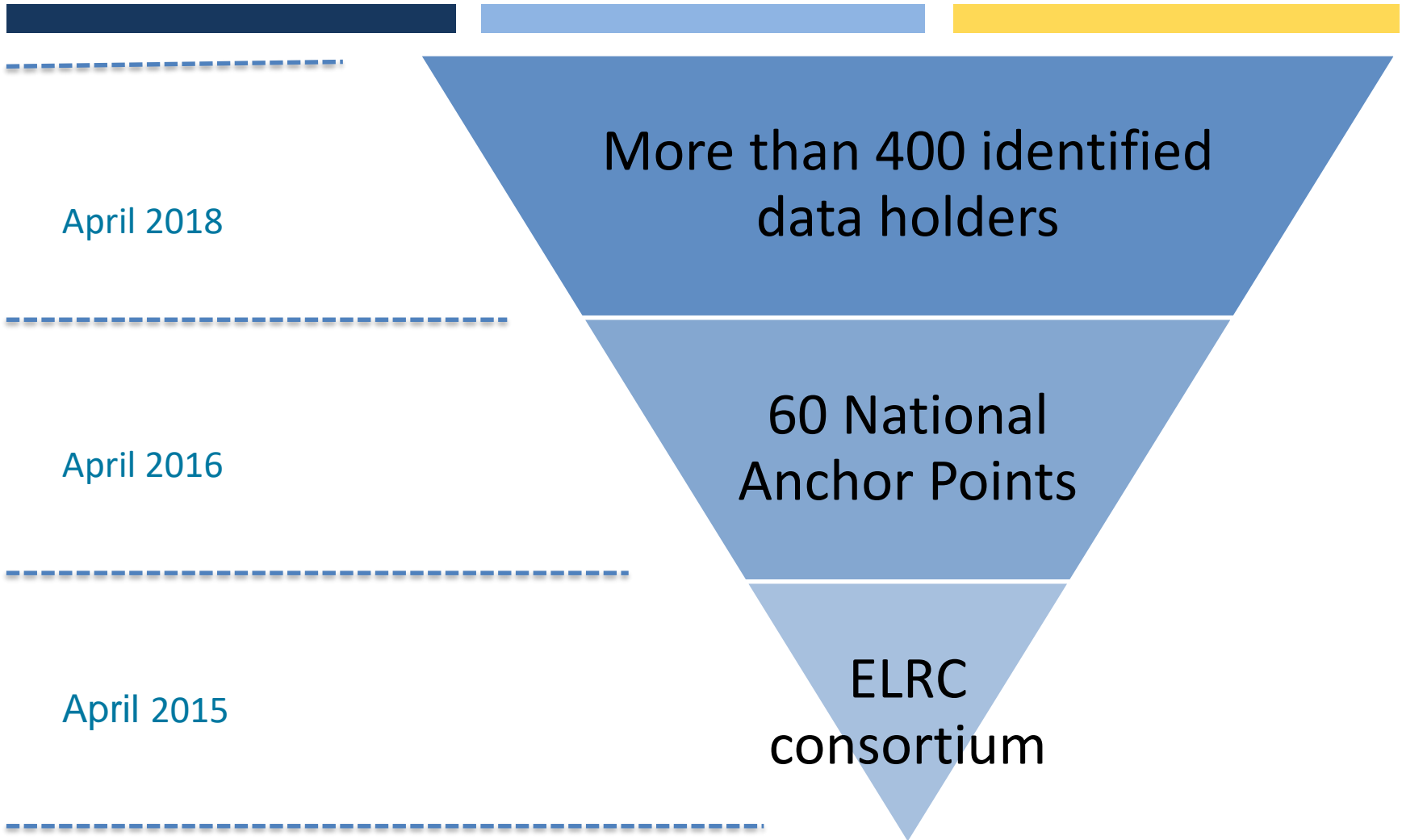




→ Your language resources make a difference!

What has been achieved so far...?

The ELRC network

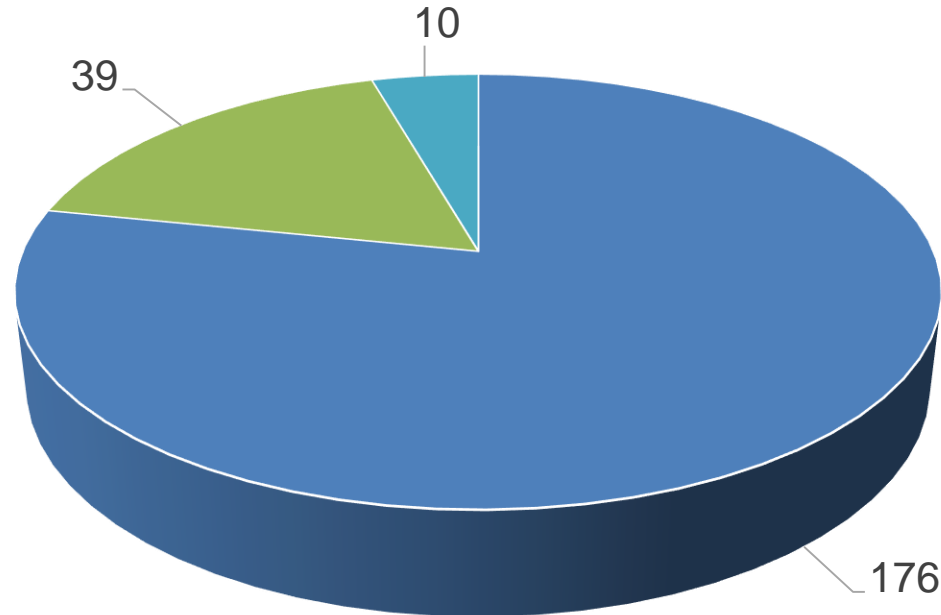


What has been achieved so far...?

ELRC Data in April 2017



A total of 225 LR provided by ELRC



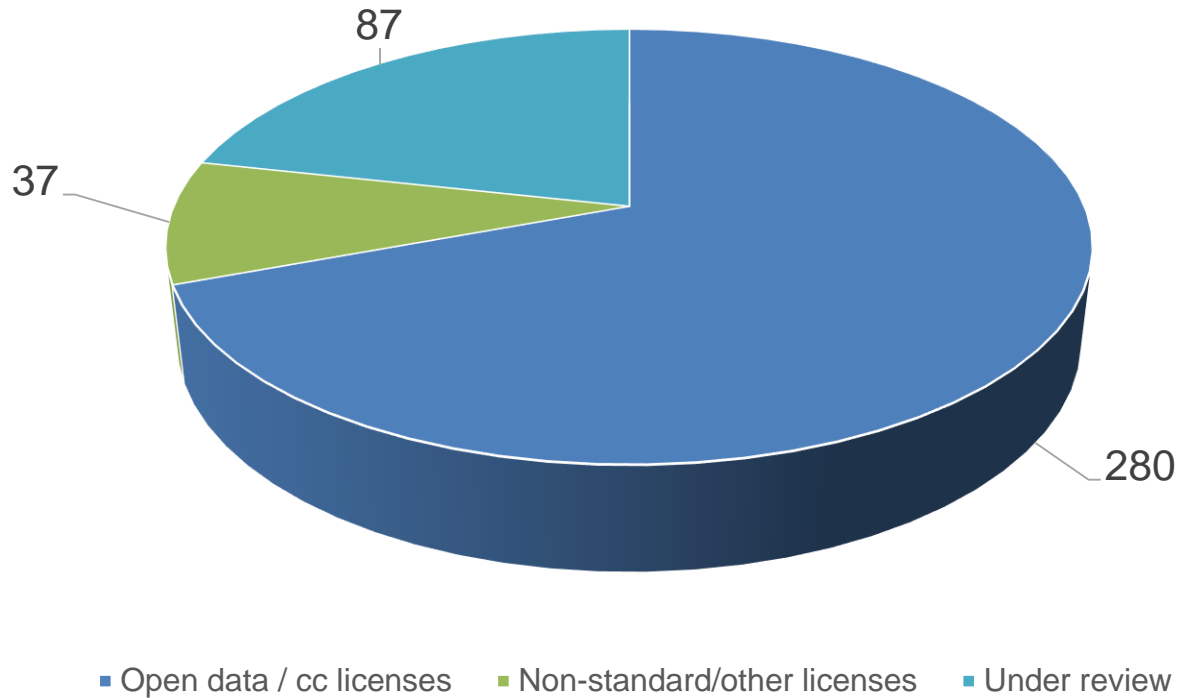
■ Open data / cc licenses ■ Non-standard/other licenses ■ Under review

What has been achieved so far...?

ELRC Data in April 2018

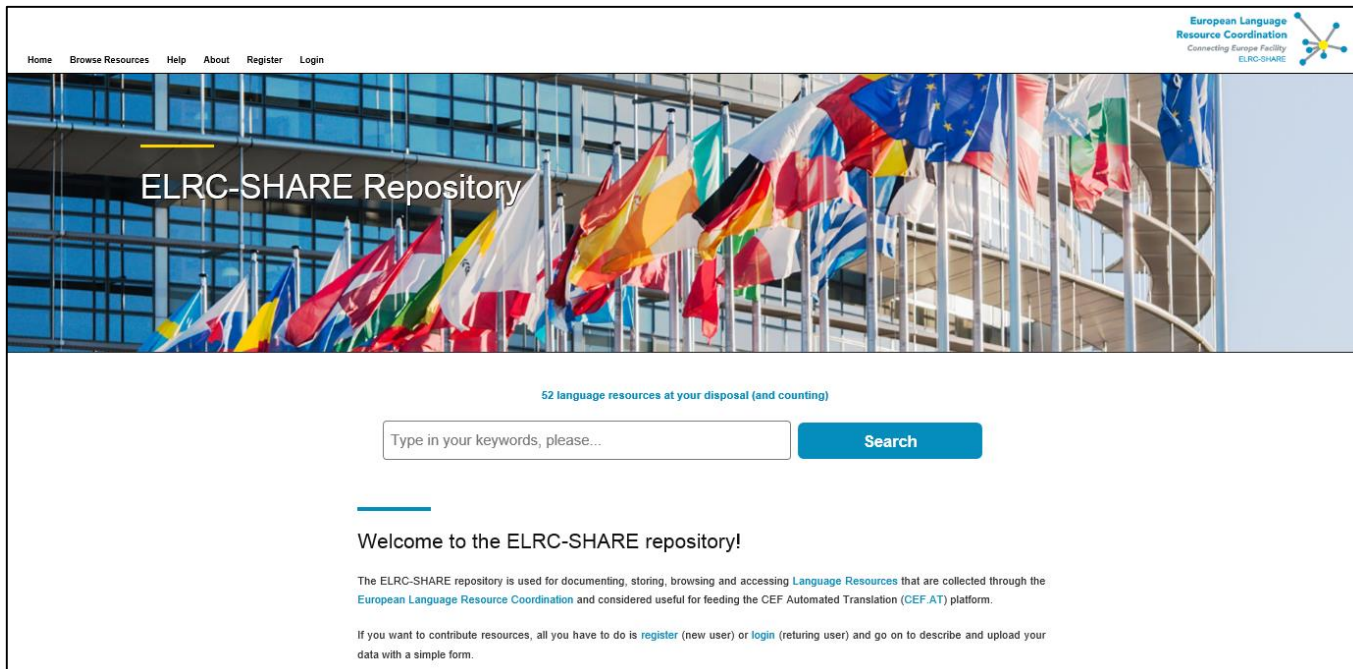


More than 400 LRs provided by ELRC



- **The ELRC-SHARE Repository**

- Access to, sharing and contribution of language resources
- Access to tools and services catalogue (upcoming)
- Visit <http://www.lr-coordination.eu/resources>



The screenshot shows the homepage of the ELRC-SHARE Repository. At the top, there is a navigation menu with links for Home, Browse Resources, Help, About, Register, and Login. The main header features the ELRC-SHARE logo and a large image of various national flags. Below the image, the text "ELRC-SHARE Repository" is displayed. A search bar is present with the placeholder text "Type in your keywords, please..." and a "Search" button. Below the search bar, the text "52 language resources at your disposal (and counting)" is shown. The main content area includes a welcome message: "Welcome to the ELRC-SHARE repository!" followed by a brief description of the repository's purpose and a link to the Language Resources section. At the bottom, there is a section for users who want to contribute resources, with links for "register" and "login".

Helpdesk for Language Resources

Helpdesk for Language Resources

We are happy to answer any questions on the technical or legal aspects related to the use, production, collection, processing, and sharing of language resources.

Please feel free to contact us through one of the following channels:

Telephone*	+33 970 440 522
Secretariat Support	+49 681 857 7552 85
Skype	ELRC Helpdesk
E-mail	help@lr-coordantion.eu

lr-coordination.eu/helpdesk

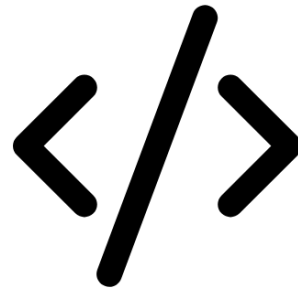


If you encounter any legal or technical issues with your language resources, ELRC and our broad network of experts can help!



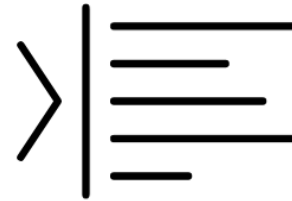
Data conversion

If your data isn't converted to the proper formats, we can help convert it



Tag removal

Does your data contain unneeded tags? We can assist in removing them!



Re-formatting

Need to re-format DOCX to XML, or PDF to WORD? Let us do it for you!



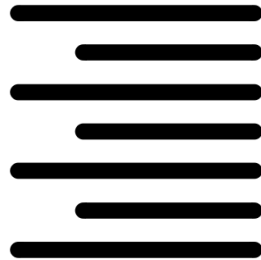
Data extraction

If your data is trapped in archives and databases, we can help extract it



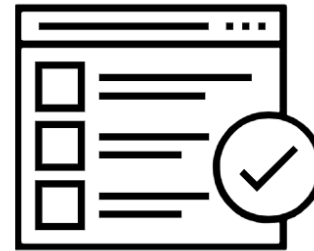
Cleaning

If your data is messy (i.e., lots of noise), we will clean it up



Alignment

Translations aren't aligned? We'll do it for you with our tools!



Metadata validation

Metadata is crucial!
We can validate and organize metadata for your team



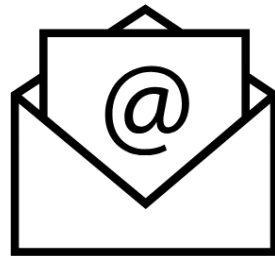
Anonymisation

Does your data contain private info?
We can help to fully anonymise



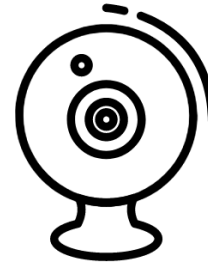
Phone

Pick up the phone and call the ELRC helpdesk



Email

Send us off an email briefly explaining your technical/legal issue



Skype

Skype us! We are online and ready to help with any questions




On-site

Our team of experts will travel directly to assist you at your own offices

ELRC consortium provides language processing services and assistance directly on-site to data contributors





We will fix your data issues and return the processed data directly to you. We can also help to improve your data management processes. Just ask!

How to request services



Submit a request for on-site assistance by filling out the form below. See a list of services [here](#).

First name *

Last name *

Institution *

Country *

Email *

Types of assistance required *

- Legal assistance
- Data processing
- Anonymisation
- Other

Description of assistance required

Submit

[Ir-coordination.eu/request-onsite-assistance](https://ir-coordination.eu/request-onsite-assistance)



Problem: In some Spanish governmental departments, archives were only available in PDF

Solution: ELRC helped provide good converters to get usable documents



Problem: Data owner needed help with anonymization, as databases contained personal info. Another need: cleaning up 'junk' data (URLs, numbers, fragments)

Solution: ELRC helped provide anonymization services and data cleaning



eTranslation TermBank

The eTranslation TermBank action aims at providing **new**
Terminological Resources (TRs) to the language resource
pipeline of CEF eTranslation to support the CEF DSIs.
More than 150 TRs covering all official languages of the EU plus
Norwegian and Icelandic are collected and will be provided on
the ELRC-SHARE repository

Case studies

Usage of eTranslation at the Swedish National Food Agency



Usage of eTranslation at the Swedish National Food Agency

General benefits: work smarter, safer, save time and money

Specific benefits for the Food Administration:
An effective tool for expert advice, risk assessment, rule development, review translations (new rules, EU or global regulations) cooperation against food fraud, global surveillance, etc.

Usage of eTranslation at the Swedish National Food Agency

Information and internal seminar

- Information on the intranet
- Prior to the seminar eTranslation manual was presented
- A live demonstration for employees at the headquarters
- Demand from the field organization (Municipalities and county administrative boards)

Lättare att förstå våra europeiska grannar – nu är eTranslation här!

Publicerad 2018-02-27 09:13

De europeiska språken är många. Nu finns en översättningstjänst som hjälper till med snabba maskinöversättningar. Användbart när du till exempel behöver sätta dig in i en text skriven på ett språk du inte behärskar. Måndagen den 5 mars har du möjlighet att lära dig mer om hur eTranslation fungerar.

Den 1 februari 2018 anslöt sig Livsmedelsverket till eTranslation, EU-kommissionens nya maskinöversättningstjänst. Tjänsten är tillgänglig för alla Livsmedelsverkets medarbetare efter registrering och inloggning.

De flesta har säkert någon gång använt sig av Google Translate. Den nya eTranslation påminner på flera sätt om denna, med skillnaden att den är bättre, säkrare, mer anpassad till myndigheters behov och ständigt utvecklas och förfinas av användare i alla EU-länder. Dessutom går det att få hela dokument översatta, inte bara enstaka stycken.

[Skriv ut](#) [Lägg till länk](#)
[Tipsa en kollega](#)
[4 Kommentarer](#)

 Connecting Europe
eTranslation

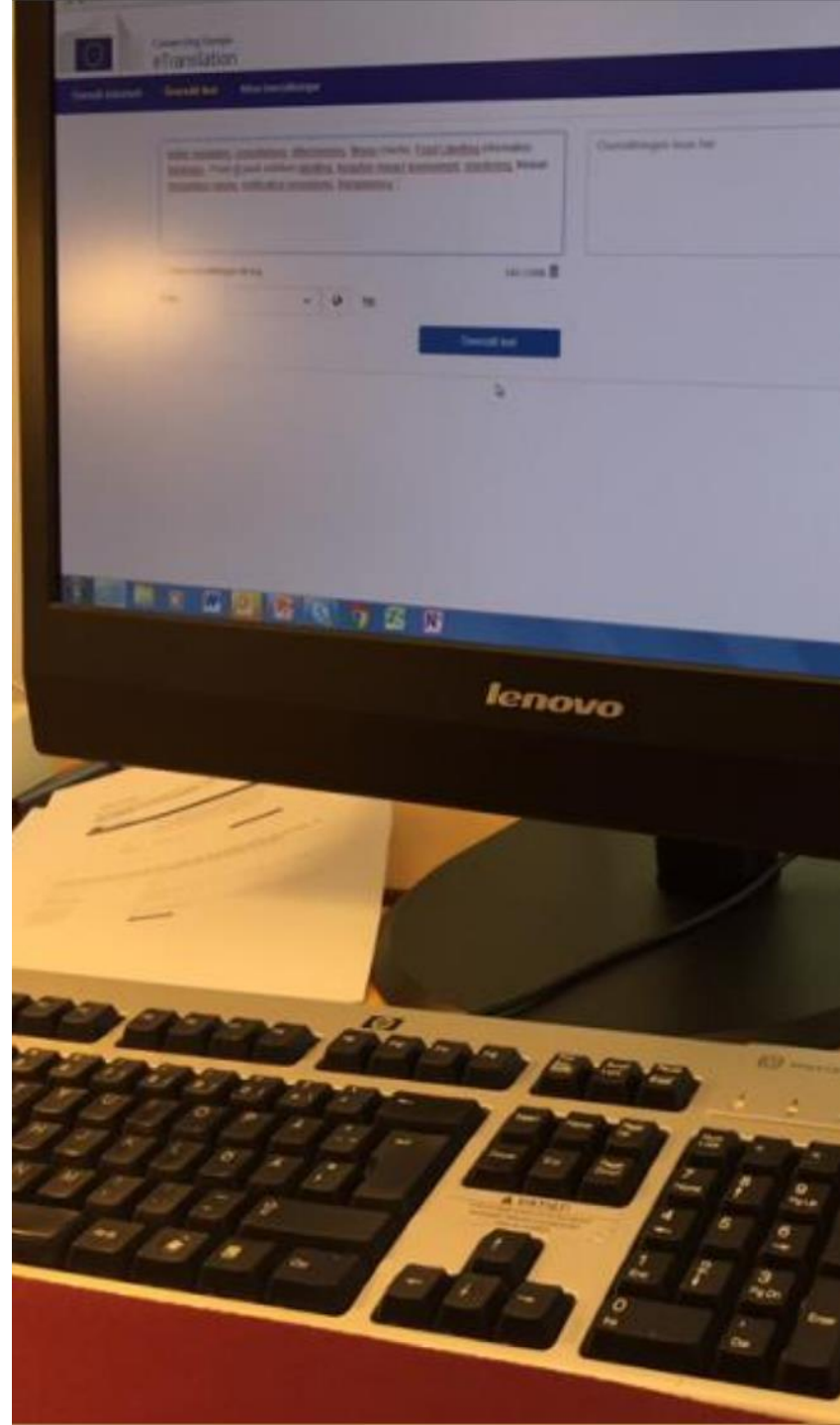
Kom igång!

Inloggning till [eTranslation](#)

Manual:
[Kom igång med eTranslation](#)

Usage of eTranslation at the Swedish National Food Agency

- Facilitate understanding of documents, news, letters, lectures / posts in other languages
- Help with emails in other languages
- Get presentation in English or another language.
- Check the ingredient list in several languages



1. Amesto Translations Ltd.

- Norway's largest commercial translation company with many public sector clients.
- Normally agreements with public agencies do not say anything about ownership or access to TM.
- The National Library has received TM from all public sector assignments since 2011, ca. 1,5 million units.
- The National Library takes responsibility for all further use.

LATVIA
TRANSLATES
WITH

HUGO.LV

- **MT service** for the Latvian public sector
- Enables **multilingual communication** in the public sector and **empowers access** to information and e-services
- **Securely** translates texts, documents, websites
- **Adapted** for the Latvian language and public sector texts
- **Integrated** into e-services and government websites
- Languages
 - Latvian-English
 - English-Latvian
 - Latvian-Russian

Developed
by



Culture Information
Systems Centre
Republic of Latvia

HUGO.LV public website

The screenshot displays the Hugo.lv public website interface. At the top left is the Hugo.lv logo. Navigation options include 'Tulkot' (highlighted), 'Atpazīt runu', and 'Izrunāt tekstu'. A user profile dropdown shows 'Jānis Lapigš' and a language dropdown shows 'Latviski'. Below the navigation, there are tabs for 'No' (Latviešu, Krievu, Angļu) and 'Uz' (Latviešu, Krievu). A document titled 'terms_and_agreement.doc' is open, with a 'Tulkot citu' button. The document content is displayed in two columns: the left column shows the English text, and the right column shows the Latvian translation. A dropdown menu is open over the 'Uz' tabs, listing 'Mani dokumenti' (8), 'Profila uzstādījumi', and 'Iziet'. A red button 'pielādēt iztulkoto' is visible on the right. The footer contains links for 'Mobilās aplikācijas', 'Rīki', 'Resursu iestatīšana', 'Meklēšana korpusos', 'Termini', 'Palīdzība', 'Atsauksmes', 'Par Projektu', and the TILDF logo.

Hugo.lv Tulkot Atpazīt runu Izrunāt tekstu Jānis Lapigš Latviski

No Latviešu Krievu Angļu Normatīvie akti

Uz Latviešu Krievu

terms_and_agreement.doc Tulkot citu

Agreement

This is a legal agreement between you, the purchaser, and the IconDrawer.com. By purchasing or downloading any royalty-free icons from our website you agree to the following:

All of the icons remain the property of IconDrawer.com. The icons can be used royalty-free by the licensee for any personal, commercial project including web design, software, application, advertising, film, video, and computer game.

The license does not permit the following uses: The icon may not be

Mani dokumenti 8

Profila uzstādījumi

Iziet

pielādēt iztulkoto

Dokuments iztulkots: ...

Nolīgums

Šī ir juridiska vienošanās starp jums, pircējam, un IconDrawer.com. Iegādājoties vai bezmaksas lejupielādes ikonu no mūsu tīmekļa vietni, jūs piekrītat šādiem nosacījumiem:

Visas ikonas paliek īpašumā IconDrawer.com. Ikonas var izmantot bezmaksas licenciātam attiecībā uz jebkuru personu, komerciālo projektu, tai skaitā Web dizaina, programmatūras, lietojumprogrammu, reklāmas, filmu, video un datorspēļu.

Licence neatļauj šādiem mērķiem: tā nav iespējams licencēt nedz

Mobilās aplikācijas Rīki Resursu iestatīšana Meklēšana korpusos Termini Palīdzība Atsauksmes Par Projektu TILDF

Machine Translation integration in the eGov portal Latvia.lv

Page content is machine translated. Latvian > English Restore original: Latvian



The Home > LIFE EVENTS > Family

Assistant

- ▶ To arrange the formalities of a child
- ▶ As electronic find personal identity number of relatives
- ▶ With the formalities relating to maintenance
- ▶ How to request a birth, death and marriage certificate or associated certificates
- ▶ Registering a pet – safety, travel and participation in fairs
- ▶ As electronically receive angling licences
- ▶ As electronically receive hunter 's seasonal card

LIFE EVENTS

- Return to Latvia
- Social services, employment
- ▶ Family
- Business
- Finance AND TAXES
- Residence AND ENVIRONMENT
- Culture, sport, tourism
- The protection of rights
- Education
- Consular services
- Vehicles
- Health
- The cross-border provision of services

- At least **35% increase of productivity and decrease of costs** for professional translations
- Opens **access to e-services** for other language communities – investors, visitors, cooperation partners, etc.
- Promotes cross- border e-services, **Pan-Baltic and Pan-European cooperation**
- **Speech tools** will facilitate e-service usage for **people with special needs**
- **Terminology database** will harmonize interinstitutional communication

Developed
by

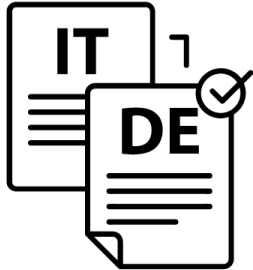


Culture Information
Systems Centre
Republic of Latvia

Integrating eTranslation for EU Council Presidency.

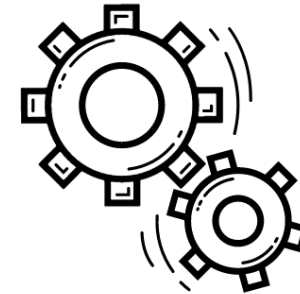


CEF eTranslation



MT systems for the **24 official EU languages** enabling translation of full documents, preserving text formatting

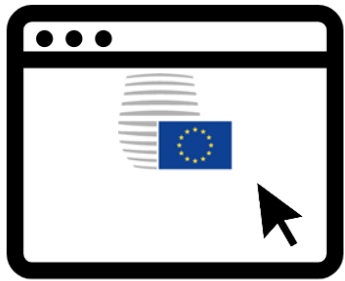
Customised MT systems



AI-powered **custom Neural MT** providing quality translation adapted for the Presidency requirements



EU PRESIDENCY TRANSLATOR



Website

For staff, widely enabling automated translation of texts, full documents, websites



Desktop

For journalists, enabling easy translation of documents, press releases, local news sites



Mobile

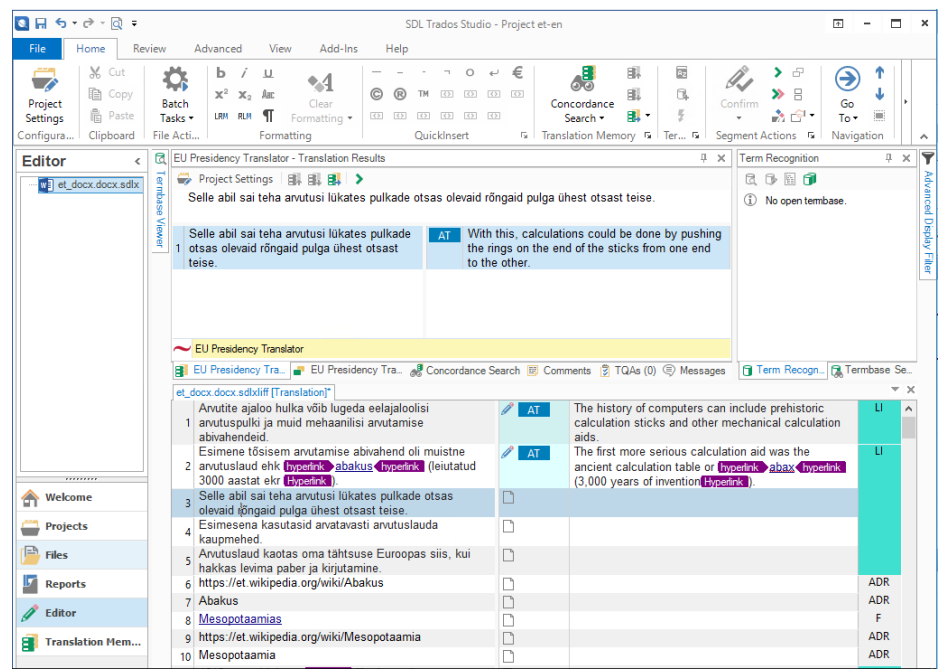
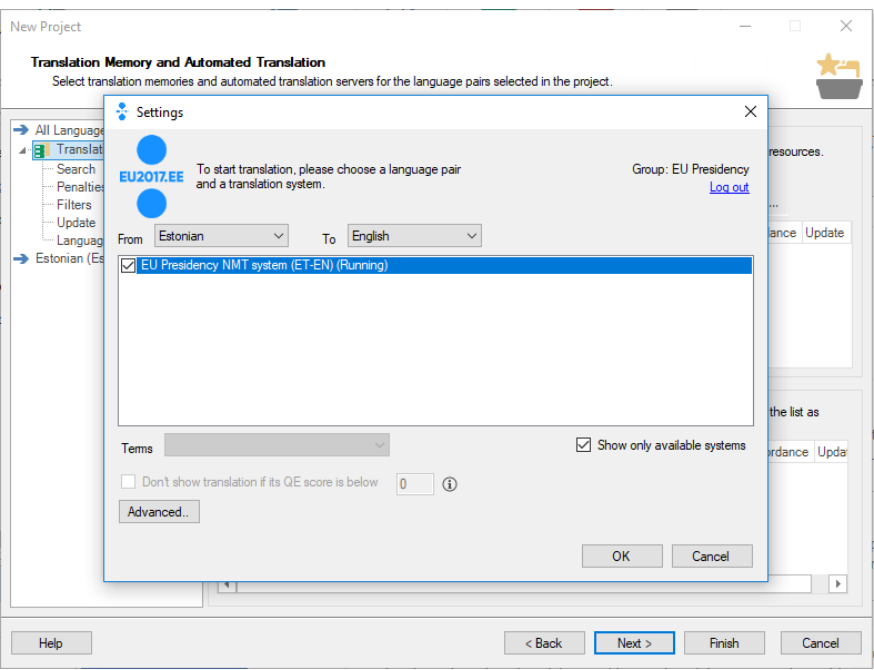
For delegates, enabling quick translation of words and phrases on the fly during meetings



Plugin

For translators, helping boost translation productivity by with MT and post-editing

PLUG-IN FOR PROFESSIONAL TRANSLATORS





Translate 2017

in cooperation with the Estonian Presidency of the Council of the EU

Menu EN EE

Translate text

Translate document

Translate website

From: English German Estonian



To: English

Translate

Eesti Euroopa Liidu Nõukogu eesistumise moto on „Ühtsus tasakaalu kaudu”. Usume, et Euroopal on piisavalt palju ühist, leidmaks edukalt lahendused kõigile väljakutsetele ning kasutamaks ära kõiki meie ees avanevaid võimalusi. Peame hõlbustama avatust nii meie majanduses kui ka ühiskonnas, kindlustades samas turvalisuse ja julgeoleku. Eesti kui eesistujariigi ülesanne on leida tasakaal liikmesriikide erinevate seisukohtade, traditsioonide ja huvide vahel, et saavutada Euroopa kodanike jaoks parim võimalik tulemus.

The motto of the Estonian presidency of the Council of the European Union is “Unity through balance”. We believe that Europe has enough in common to successfully find solutions to all challenges and to exploit all the capabilities that we face. We need to facilitate openness in both our economies and in society, while securing security and security. The task of Estonia as the presidency is to find a balance between the various views, traditions and interests of member states in order to achieve the best possible outcome for European citizens.

Powered by **eTranslation**, the EU's automated translation service

Developed by **Tilde**

translate2017.eu

EU2017.EE



Bulgarian Presidency of the Council of the European Union



- Translate text
- Translate document
- Translate website

From: English German French



To: German French Bulgarian

Translate

Enter the text you want to translate.

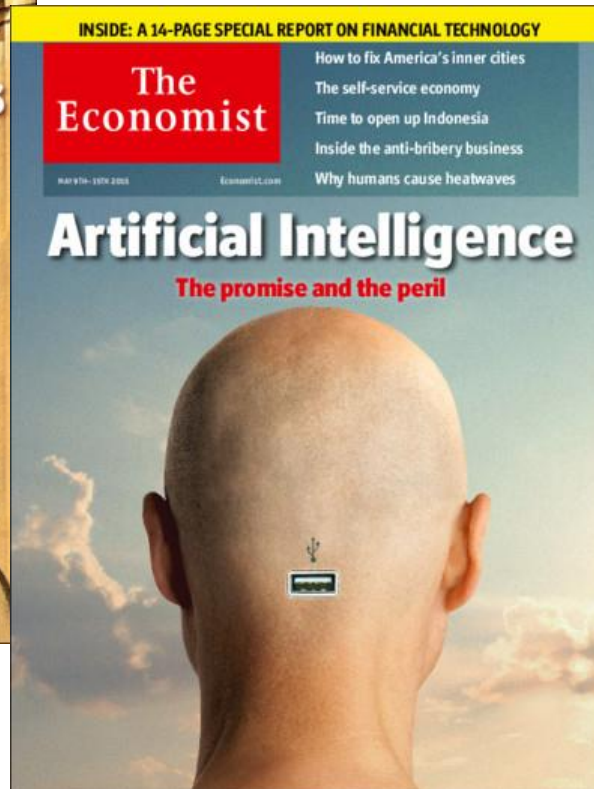
Machine translation results help to understand the meaning of a source text, but do not equal translation by a human.

EU COUNCIL PRESIDENCY TRANSLATOR



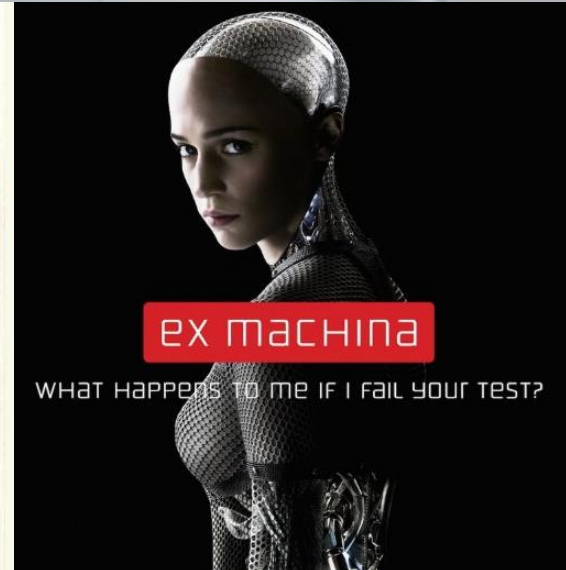
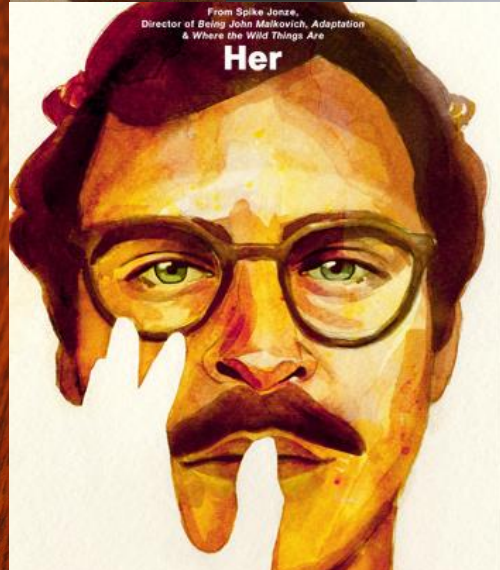


Next challenge: to make Artificial Intelligence multilingual





Hi, I'm Cortana.



European Parliament Resolution on Language Equality in the Digital Age

(Sept 11, 2018)



Language equality in the digital age

European Parliament resolution of 11 September 2018

whereas

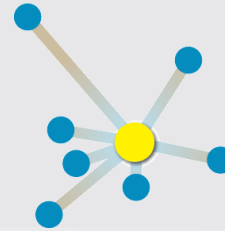
- an **enormous amount of data** is expressed in human languages
- the management of LTs could enable a wide range of **innovative IT products and services** in industry, commerce, government, research, public services and administration, reducing natural barriers and market costs...

stresses

- **the importance of text** and data mining techniques for the development of language technologies
- the need to strengthen **collaboration between industry and data owners**
- the need to ensure a more **open and interoperable use** and **collection of language resources**

Ačiū labai!

www.lr-coordination.eu
info@lr-coordination.eu



**European Language
Resource Coordination**
Connecting Europe Facility

