



# The CEF eTranslation platform @ work

Claus Thornby Larsen  
Language Technology Coordinator  
Generaldirektoratet for Oversættelse  
Europa-Kommissionen

## MT@EC

- Launched June 2013
- Legalese
- Statistical (Moses)

## eTranslation

- Launched July 2017 (webservice for snippets), Nov. 2017 (web page for documents)
- Cloud based
- Neural engines

## CEF.AT

- More NLPs (transliteration, named entity recognition...)
- Generic services
- Projects (ELRC, market research)

- Available for:
  - individuals (submit documents through a web page)
  - machine-to-machine use
- Users:
  - Digital Service Infrastructures (EESI, ODR, Open Data Portal, Europeana, etc.)
  - System suppliers (EURLex, N-Lex, Internal Market Information system...)
  - Individuals in public administrations
- Benefits:
  - Increase speed and productivity
  - Reduce costs
  - Facilitate information exchange

# Human-to-machine services

Connecting Europe  
eTranslation

External  
About | Help

Translate documents   Translate text   My translation requests   My settings   English   Logout

Drop files to upload (or click)

Supported formats:

**From \*** E ▾

**To \***

BG	CS	DA	DE	EL	ES	ET	FI	FR	GA	HR	HU	IS	IT	LT	LV
MT	NB	NL	PL	PT	RO	SK	SL	SV							

**Domain** Cutting edge (default) ▾

**Output format** Same as source TMX XLIFF

**Translate document**

Submit your document...



EUROPEAN COMMISSION

MEMO

Brussels, 6 May 2013

## **Smarter rules for safer food: Commission proposes landmark package to modernise, simplify and strengthen the agri-food chain in Europe**

### **1) What is "from farm-to-fork" policy?**

The EU "from farm-to-fork" policy covers legislation to **protect health** by ensuring that the food we produce and import that ends up on our plates is safe to eat. This body of legislation is a complex, integrated system of rules which takes care of all the links in the food chain from: the protection of plants from pests, the protection of animals and humans from animal diseases (and from such diseases which are transmissible to humans) and the protection of human health from all the hazards that might arise along the agri-food chain. It also includes all those rules which allow the Single Market to function smoothly in this area and provides rules for the protection of consumers.

Official controls are carried out at various stages in the food production chain by EU Member States to ensure that the legislation is being correctly applied.

Get your translation  
back by e-mail.



Règles plus intelligentes pour des denrées alimentaires «de la fourche à la fourchette»: Paquet décisif pour moderniser, de simplifier et renforcer le secteur agroalimentaire en Europe

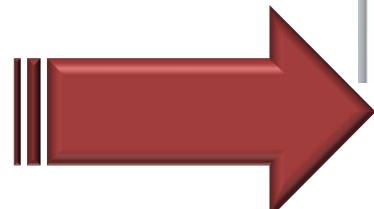
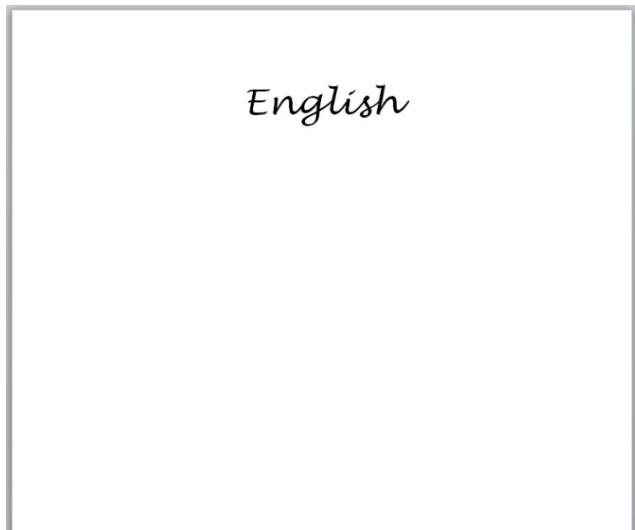
**1) Ce qui est «de la fourche à la fourchette»**

L'UE «de la fourche à la fourchette» a pour objectif de faire de la santé en faisant en sorte que l'agriculture et l'alimentation sont au cœur de la politique finale sur notre agriculture et notre alimentation. Le système législatif est un système de maillons de la chaîne alimentaire qui protège les citoyens contre les dangers nuisibles, la protection de l'environnement et la sécurité. Partant de ces malades, l'UE a mis en place une législation qui protège tous les dangers et les risques pour la santé humaine. La législation comprend également la protection des animaux et la sécurité des consommateurs.

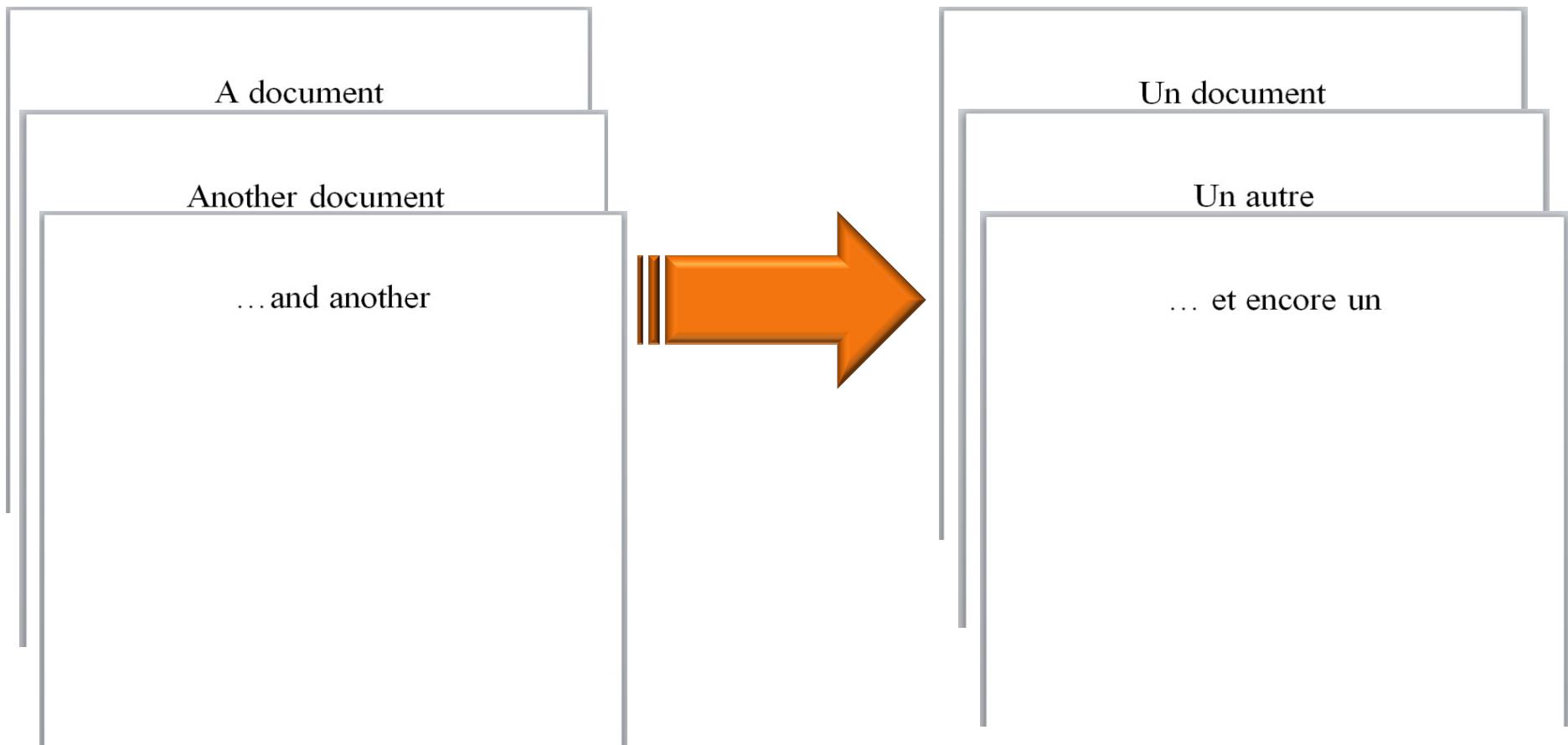


Keeps your  
formatting!

One original,  
many translations...



## Many documents at once...



## eTranslation protects your privacy

- All documents deleted after 24 hours or after delivery (on demand)
- IPR rights transferred  
**(Your translation belongs to YOU!)**

No one

looking

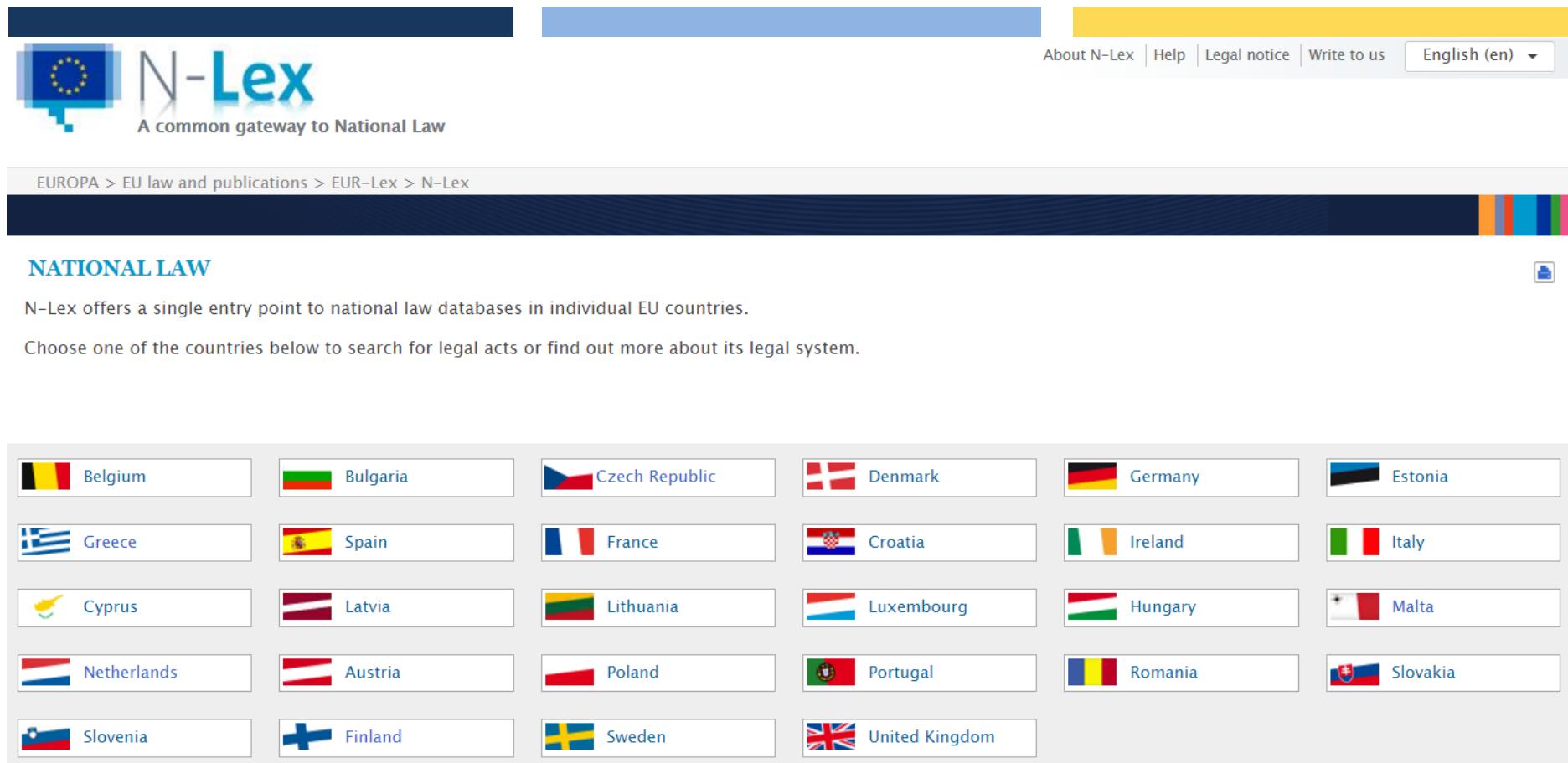
at your translations!

# Machine-to-machine services

## eTranslation on the N-Lex portal



European Commission

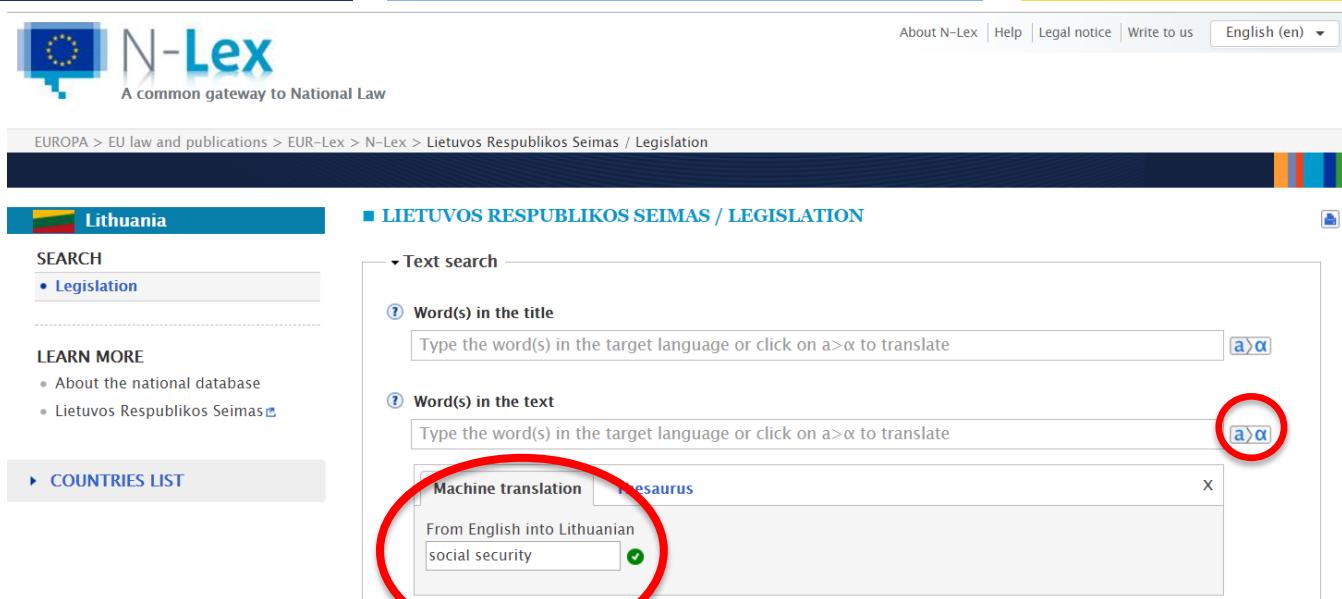


The screenshot shows the N-Lex portal interface. At the top, there is a navigation bar with links to "About N-Lex", "Help", "Legal notice", "Write to us", and a language selector set to "English (en)". Below the navigation bar, the N-Lex logo is displayed, followed by the text "A common gateway to National Law". A breadcrumb navigation path is shown: "EUROPA > EU law and publications > EUR-Lex > N-Lex". On the right side of the header, there is a small graphic of colored bars. The main content area is titled "NATIONAL LAW" and contains the following text: "N-Lex offers a single entry point to national law databases in individual EU countries. Choose one of the countries below to search for legal acts or find out more about its legal system." Below this text, there is a grid of 28 boxes, each representing a European country with its flag and name: Belgium, Bulgaria, Czech Republic, Denmark, Germany, Estonia, Greece, Spain, France, Croatia, Ireland, Italy, Cyprus, Latvia, Lithuania, Luxembourg, Hungary, Malta, Netherlands, Austria, Poland, Portugal, Romania, Slovakia, Slovenia, Finland, Sweden, and United Kingdom.

# Machine-to-machine service eTranslation on the N-Lex portal



European Commission



The screenshot shows the N-Lex portal interface for Lithuania. On the left, there's a sidebar with a search bar set to 'Legislation' and links for 'About the national database' and 'Lietuvos Respublikos Seimas'. Below that is a 'COUNTRIES LIST' button. The main content area is titled 'LIETUVOS RESPUBLIKOS SEIMAS / LEGISLATION' and includes a 'Text search' section. This section has two input fields: 'Word(s) in the title' and 'Word(s) in the text'. Both fields have a placeholder 'Type the word(s) in the target language or click on a>α to translate' and a blue 'a>α' button. A red circle highlights this button. Below these fields is a dropdown menu with 'Machine translation' selected, showing a preview of 'From English into Lithuanian' with the text 'social security' and a green checkmark. Another red circle highlights the 'a>α' button in this dropdown.

## ① Word(s) in the text

socialinė apsauga ←

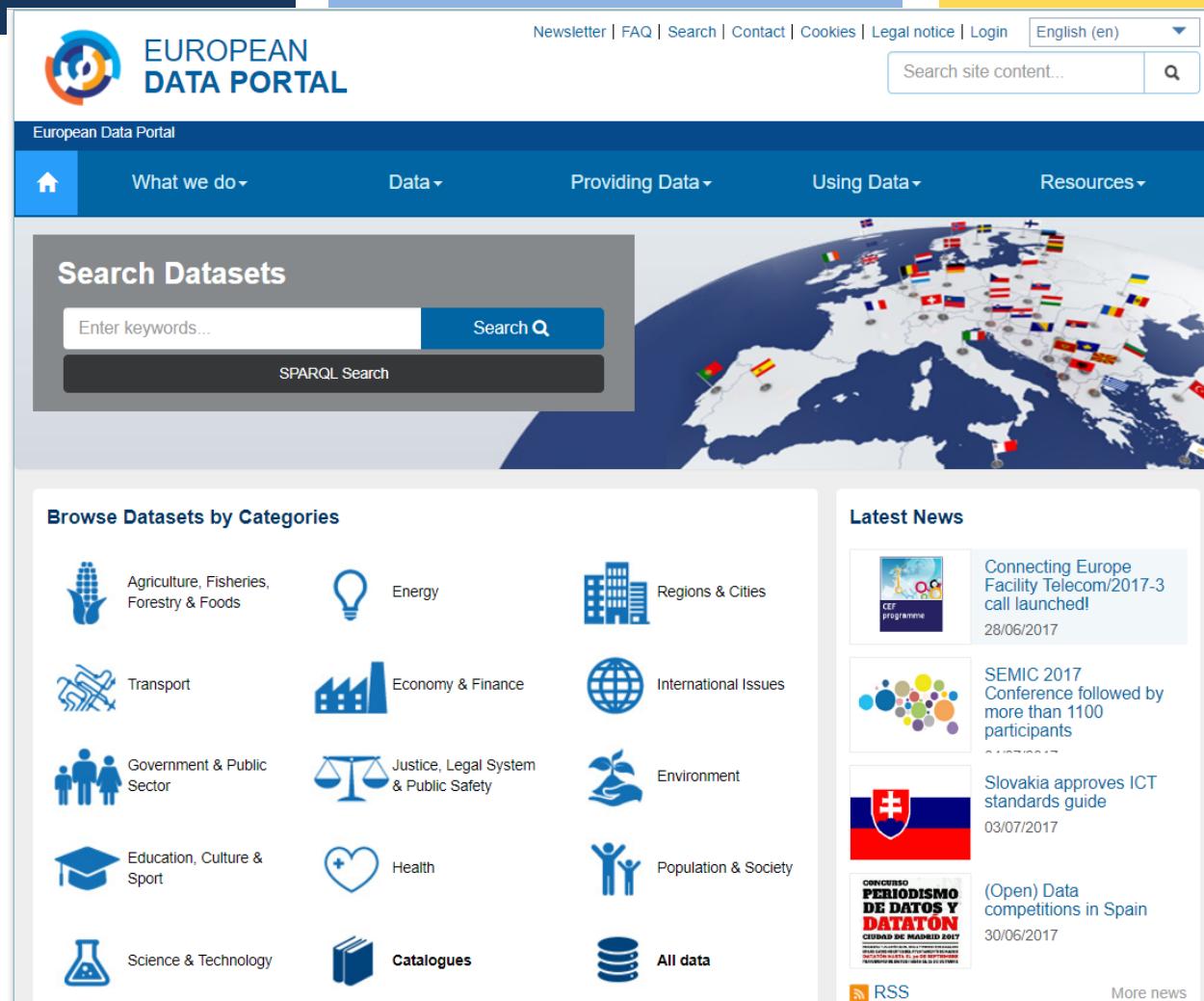
a>α

# Machine-to-machine service

## eTranslation on the European Data Portal



European Commission



The screenshot shows the European Data Portal homepage. At the top, there is a navigation bar with links to Newsletter, FAQ, Search, Contact, Cookies, Legal notice, and Login, along with a language selector set to English (en). Below the navigation is a search bar with the placeholder "Search site content..." and a magnifying glass icon.

The main content area features a large map of Europe with national flags overlaid, representing data providers. To the left of the map is a search interface with a "Search Datasets" heading, a keyword input field, a "Search" button, and a "SPARQL Search" button. Below this is a section titled "Browse Datasets by Categories" containing 12 icons and labels:

- Agriculture, Fisheries, Forestry & Foods
- Energy
- Regions & Cities
- Transport
- Economy & Finance
- International Issues
- Government & Public Sector
- Justice, Legal System & Public Safety
- Environment
- Education, Culture & Sport
- Health
- Population & Society
- Science & Technology
- Catalogues
- All data

To the right of the categories is a "Latest News" sidebar with four news items:

- Connecting Europe Facility Telecom/2017-3 call launched! (28/06/2017)
- SEMIC 2017 Conference followed by more than 1100 participants (04/07/2017)
- Slovakia approves ICT standards guide (03/07/2017)
- (Open) Data competitions in Spain (30/06/2017)

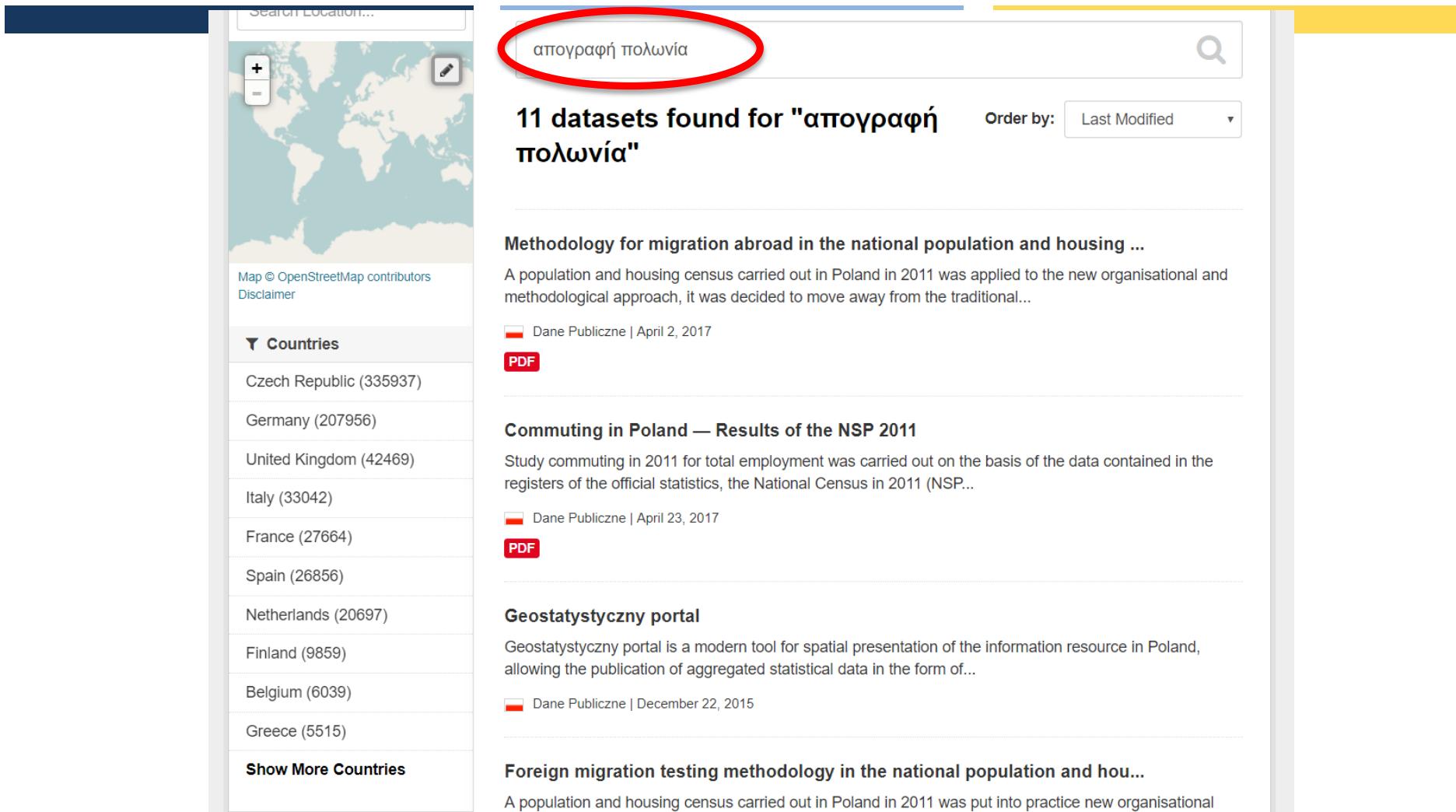
At the bottom right of the sidebar is an "RSS" feed icon.

# Machine-to-machine service

## eTranslation on the European Data Portal



European Commission



The screenshot shows the European Data Portal's search interface. A red circle highlights the search bar containing the Greek text "απογραφή πολωνία". Below the search bar, the results are displayed:

**11 datasets found for "απογραφή πολωνία"**

Order by: Last Modified ▾

**Methodology for migration abroad in the national population and housing ...**  
A population and housing census carried out in Poland in 2011 was applied to the new organisational and methodological approach, it was decided to move away from the traditional...  
Dane Publiczne | April 2, 2017  
[PDF](#)

**Commuting in Poland — Results of the NSP 2011**  
Study commuting in 2011 for total employment was carried out on the basis of the data contained in the registers of the official statistics, the National Census in 2011 (NSP)...  
Dane Publiczne | April 23, 2017  
[PDF](#)

**Geostatystyczny portal**  
Geostatystyczny portal is a modern tool for spatial presentation of the information resource in Poland, allowing the publication of aggregated statistical data in the form of...  
Dane Publiczne | December 22, 2015

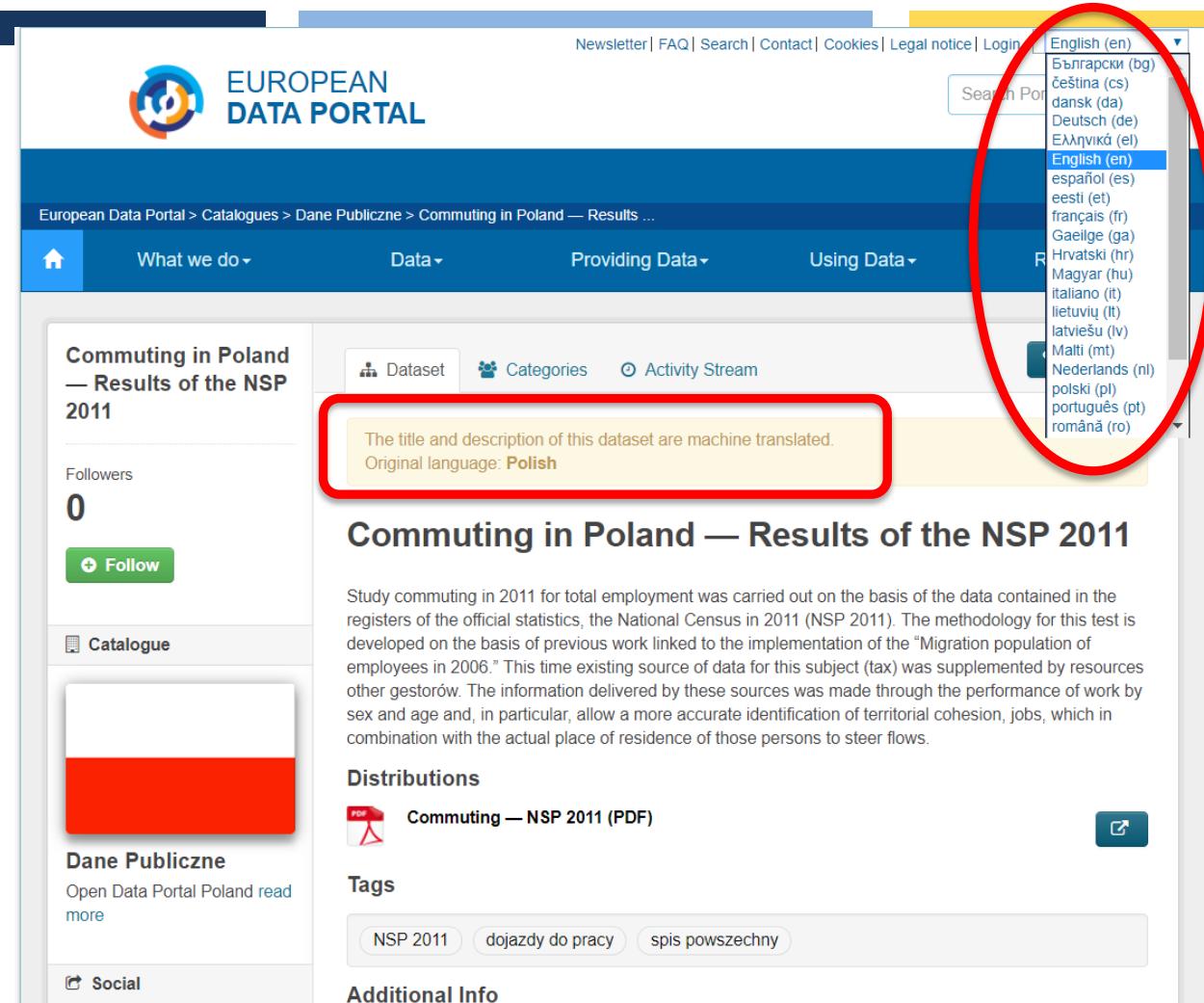
**Foreign migration testing methodology in the national population and hou...**  
A population and housing census carried out in Poland in 2011 was put into practice new organisational

# Machine-to-machine service

## eTranslation on the European Data Portal



European Commission



The screenshot shows the European Data Portal interface. At the top right, a language dropdown menu is open, listing various languages with English (en) selected. A red circle highlights this dropdown. Below it, a message box states: "The title and description of this dataset are machine translated. Original language: Polish". This message is also highlighted with a red box. The main content area displays a dataset titled "Commuting in Poland — Results of the NSP 2011". The dataset summary includes a brief description of the study, distributions (a PDF file), and tags (NSP 2011, dojazdy do pracy, spis powszechny). On the left sidebar, there are sections for "Followers" (0), "Catalogue" (Dane Publiczne), and "Social".

# How to Connect to eTranslation

- Open to DSIs and public administrations
  - Contact [CEF-AT@ec.europa.eu](mailto:CEF-AT@ec.europa.eu) with your request and use case
  - We will provide the technical documentation on how to connect
    - SOAP request or RESTful interface
  - Contact us for credentials
  - Adapt your service to multilingual use!

# Behind the scenes: how it works

- Statistical Machine Translation
  - MT@EC built only on "EU translations" (Euramis database)
  - Covers all 24 languages
  - Euramis (and MT@EC) cover EU policies, subjects and language but limited everyday language
  - Moses-based engines
  - Performs best when trained on **large volumes** of text pairs (source-translated) in **specific domains**

## What is NMT?

- Machine learning: artificial neural networks trained on existing translations
- The computer devises its own rules on how to translate
- Radical departure from the phrase-based SMT approaches

## Why is it important?

- Translations read better: more fluent and grammatical
- Better able to fill in gaps in data used for training
- Better for highly inflected languages (e.g. German, Hungarian)

## This is where the field is headed

- Will SMT become obsolete?

Hungarian Original: **Hatalmas bevételt titkolnak el a vállalkozások**

Az online kassza önmagában nem hozhat, és nem is hozott érdemi változást az eltitkolt bevételek csökkentésében – legalábbis erre a következtetésre jutottak a lapunk által megkérdezett szakértők a Nemzeti Adó- és Vámhivatal (NAV) új akciósorozata, az úgynevezett adótraffipax sikere láttán.

**MT@EC: The huge revenue being hidden**

Online **pot may not in itself, and not the substantive change to reduce hidden revenue at least this was the conclusion reached** by the experts consulted by the debates of the National Tax and Customs Administration (NTCA), the so-called **akciósorozata adótraffipax** success.

**eTranslation: Huge revenue will be hidden for businesses**

The online retention cannot, and did not, make a significant difference in the reduction of the hidden revenue — at **least this conclusion was concluded** by the experts consulted by us and the new action of the National Tax and Customs Administration (NAV) — **the success of the so-called “tax force”**.

From **Magyar Nemzet**, 17 July (<https://mno.hu/gazdasag/hatalmas-bevetelt-titkolnak-el-a-vallalkozasok-2408682>)

## eTranslation of a news article

Source: DA

Skattesag. Jesper 'Kasi' Nielsen vandt i går i Højesteret et opgør med Skat om et skattesmæk på 245 millioner kroner. Sagen går tilbage til forretningsmandens samarbejde med smykkeselskabet Pandora.

Det handler om, hvorvidt han var berettiget til at blive fritaget for skat, da han i december 2009 overførte værdier til et nyt selskab i forbindelse med samarbejdet.

Skat krævede 245 millioner kroner. Det er skat af en gevinst, som Jesper 'Kasi' Nielsen fik, da Pandora siden blev børsnoteret i 2010, og han blev købt ud af selskabet.

[Jyllands-Posten, 5.10.2018]

Target: EN (NMT)

Tax case. Mr Jesper 'Kasi' Nielsen took off, in the Supreme Court, a report on a tax reduction of DKK 245 million. The case goes back to the business of the business of the jewellery company pandora.

It is about whether he was entitled to be exempt from tax, as in December 2009 he transferred the values to a new company in the context of the cooperation.

SKAT required DKK 245 million. It is tax on a gain that Jesper 'Kasi' Nielsen was granted since pandora later was listed in 2010 and was bought out by the company.

# eTranslation of a news article on technology domain

Source: DA

Der er indgået en endelig aftale i Europa-Parlamentet og ministerrådet om en afskaffelse af roamingafgifter i EU.

Afskaffelsen træder i kraft 15. juni 2017.  
[...]

- Det har været en lang kamp, der har stået på siden 2013. Dette er en sejr for forbrugerne, at de nu kan bruge data i alle EU-lande, uden at det skal koste dem ekstra, siger han.

[nyheder.tv2.dk, 1.2.2017]

Target: EN (NMT)

A final agreement has been reached in the European Parliament and the Council of Ministers on the abolition of roaming charges in the EU.

Its abolition will enter into force on 15 June 2017.

[...]

— It has been a long battle that has been active since 2013, and consumers are now able to use data in all EU countries without extra cost, he says.

Source: DA

Louisiana har en god kontakt til fransk musikliv, som hvert år kaster interessante kammerkoncerter af sig. Torsdag gælder det strygekvartetten Modigliani, som har den sjældne kvalitet både at kunne spille raffineret og klangskønt – og give den gas, som om det gjaldt livet. De begynder dog i den lette ende med et divertimento af Mozart, der hører til blandt de mest spiselige godbidder fra komponistens hånd

[Politiken, 5.10.2018]

Target: EN (NMT)

Louisiana have a good contact with French music lives that shout an interesting club concerts each year. On Thursday, the candiani, which has the rare quality, will both be able to play both a refined and a and-long way, and to provide the gas as if it were to do so. They do, however, begin in the easy end with a divertimento of Mozart which is one of the most edible treats from the part of the composer

Source: EN

THE EUROPEAN COMMISSION,  
Having regard to the Treaty on the Functioning  
of the European Union,  
[...]

HAS ADOPTED THIS DECISION:

Article 1

The total sum of greenhouse gas emissions  
covered by Decision No 406/2009/EC for each  
Member State for the year 2014 arising from  
the corrected inventory data upon completion  
of the comprehensive review referred to in  
Article 19(1) of Regulation (EU) No 525/2013 is  
set out in the Annex to this Decision.

[...]

Target: DA (NMT)

EUROPA-KOMMISSIONEN HAR —  
Under henvisning til traktaten om Den  
Europæiske Unions funktionsmåde,  
[...]

VEDTAGET FØLGENDE BESLUTNING:

Artikel 1

Den samlede mængde drivhusgasemissioner,  
der er omfattet af beslutning nr. 406/2009/EF,  
i henhold til de korrigerede opgørelsesdata for  
2014 for de enkelte medlemsstater efter  
afslutningen af den samlede revision, som er  
omhandlet i artikel 19, stk. 1, i forordning (EU)  
nr. 525/2013, findes i bilaget til denne  
afgørelse.

[...]

Source: EN

THE EUROPEAN COMMISSION,  
Having regard to the Treaty on the Functioning  
of the European Union,  
[...]

HAS ADOPTED THIS DECISION:

Article 1

The total sum of greenhouse gas emissions  
covered by Decision No 406/2009/EC for each  
Member State for the year 2014 arising from  
the corrected inventory data upon completion  
of the comprehensive review referred to in  
Article 19(1) of Regulation (EU) No 525/2013 is  
set out in the Annex to this Decision.

[...]

Target: DA – (published translation)

EUROPA-KOMMISSIONEN HAR —  
Under henvisning til traktaten om Den  
Europæiske Unions funktionsmåde,  
[...]

VEDTAGET FØLGENDE  
**BESLUTNINGAFGØRELSE:**

*Artikel 1*

Den samlede mængde drivhusgasemissioner,  
der er omfattet af beslutningafgørelse nr.  
406/2009/EF, i henhold til de korrigerede  
opgørelsesdata for 2014 for de enkelte  
medlemsstater efter afslutningen af den  
samlede revision, som er omhandlet i artikel  
19, stk. 1, i forordning (EU) nr. 525/2013,  
findes i bilaget til denne afgørelse.  
[...]

# Language resources: the key to success



European Commission



- More data for all languages
- Better lexical coverage for all languages
- Machine translation adapted to your domain/subject matter

# Future Improvements

Requested by users such as EDP, e-Justice, Europeana, EESSI, ODR...

- Speed
- New languages (AR, RU, TK, ZH)
- More formats (PDF output, JSON)
- Named-entity recognition
- Transliteration

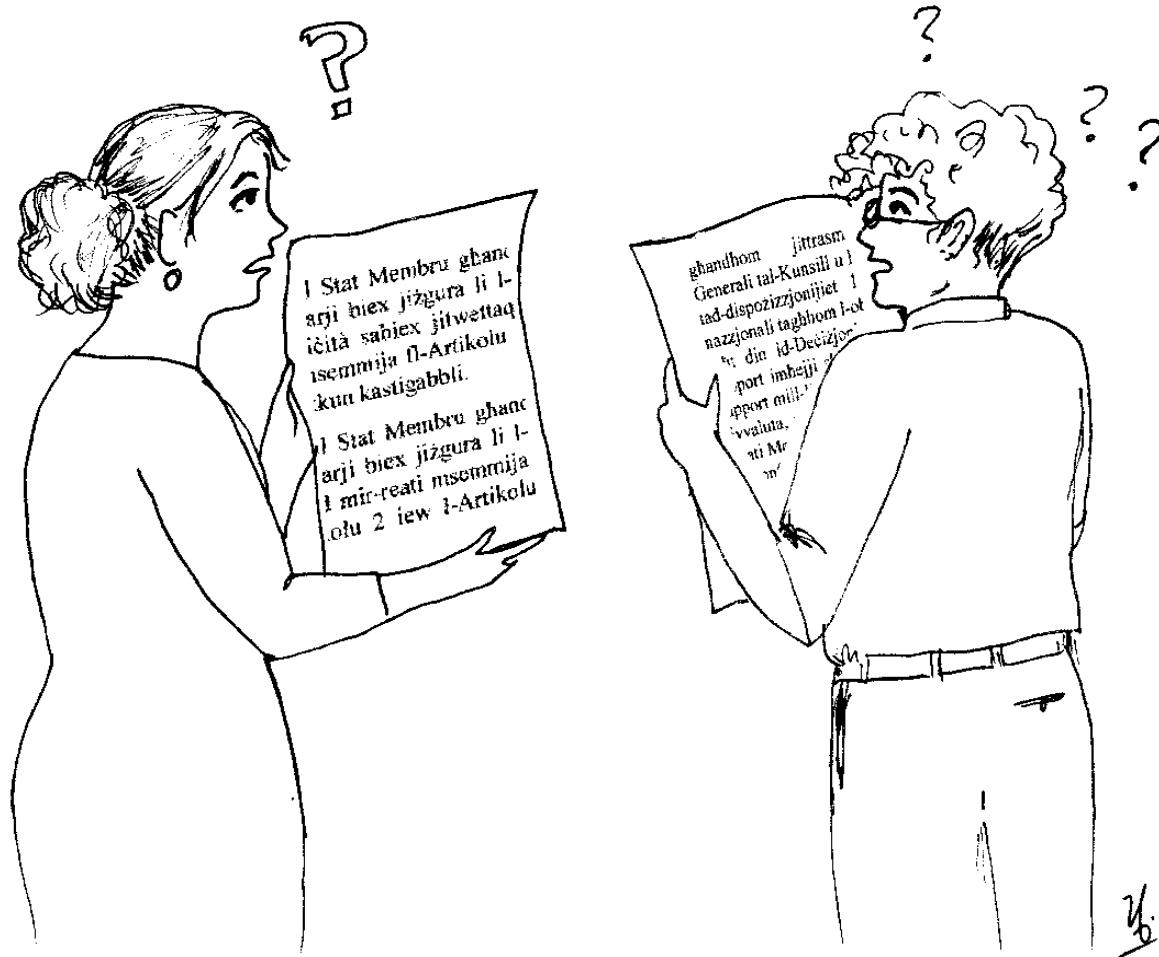
## CEF.AT brings...



European Commission

- Reliable and trustworthy translation for EU and National Public Administrations
- Support for languages with fewer speakers
- Opportunities for private sector, through grants and generic services
- A higher profile for language technologies, thus fostering demand
- Public availability of data collected by ELRC

# Solve your language problems!



Connecting Europe  
eTranslation



Thank you!