

Country Profile: The Netherlands

Jan Odijk & Carole Tiberius

Language Resource Board, Amsterdam

2019-03-27

- The Dutch Language
- Country Report
- Identified Data

➤ The Dutch Language

- Country Report
- Identified Data

- App.24 million native speakers:
 - 16 million in the Netherlands
 - 6.5 million in Belgium (Flanders)
 - 400,000 in Suriname.
- Dutch is one of the 40 most spoken languages in the world
- Official language in the Netherlands and in Flanders
- One of the official languages in
 - Belgium, Surinam, Aruba, Curacao and Sint-Maarten
- also spoken
 - in the EU in France and Germany,
 - Outside the EU in Brazil, Canada, Indonesia (Java and Bali), South Africa, and the United States.

- **Strong online:**
 - it is one of the 10 most important languages on the Internet and social media.
 - 8th most used language on Twitter
 - 9th language on Wikipedia.
 - **.nl** is the most well-known extension of the Netherlands:
 - > 5.6 million **.nl** sites registered
 - 5th most popular country extension in the world.
 - A lot of Dutch also on most **.be** websites,

- The Dutch Language
- **Country Report**
- Identified Data

- Not centralised (not even within a ministry)
- A number of ministries and executing bodies have their own translation department, e.g.
 - Ministry of Foreign Affairs
 - Ministry of Justice and Security, especially for interpreting (Foreigners chain, Ministry of Justice and Security)
 - Ministry of Defense
 - National Police
 - General information and Security Service (AIVD (<https://www.aivd.nl/>)
 - Military Information and Security Service (MIVD) (<https://www.defensie.nl/organisatie/bestuurstaf/eenheden/mivd>)
 - Social Security Bank (Sociale VerzekeringsBank, SVB)
 - Authority for Consumers and Market (ACM)

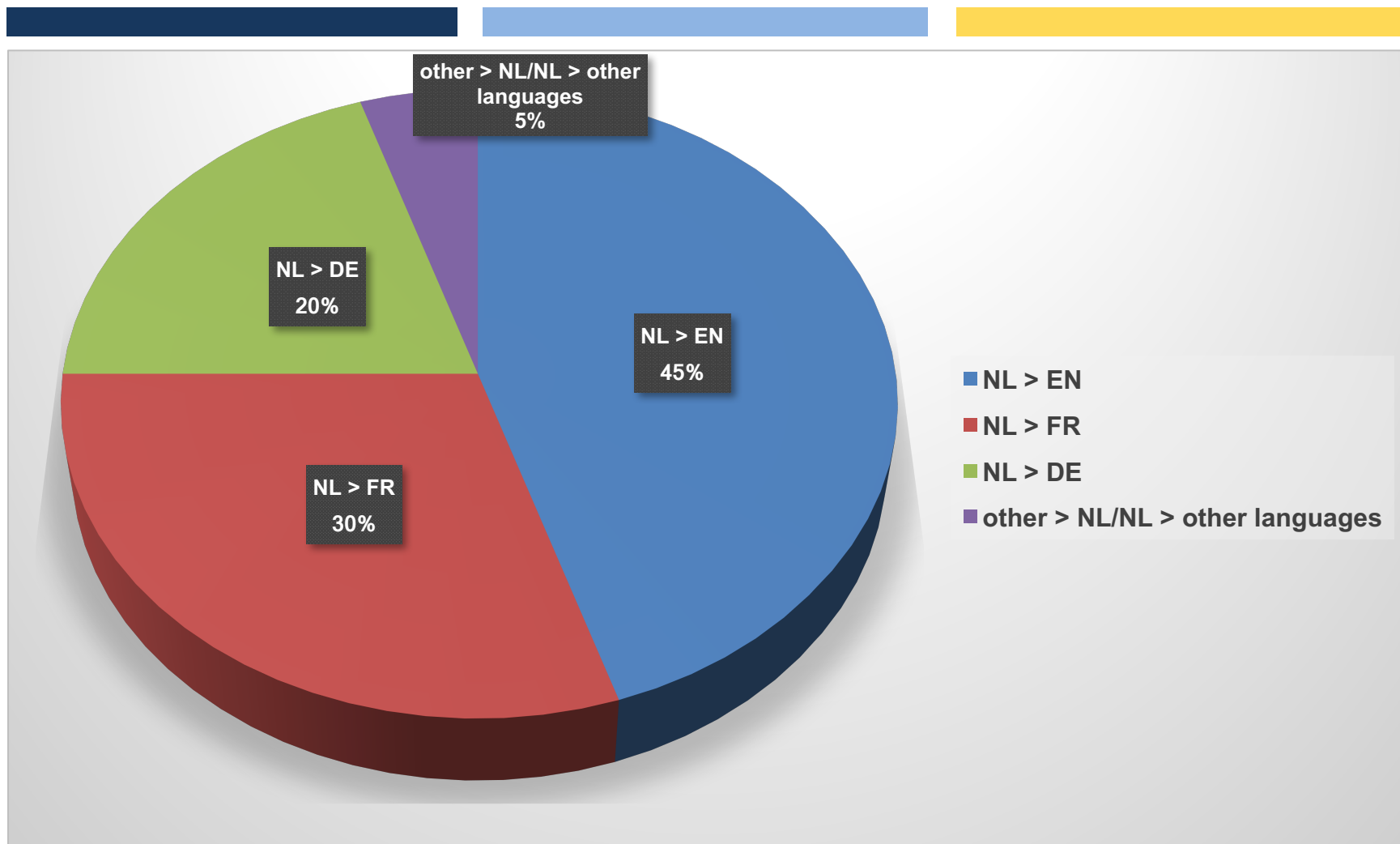
- Most translation is outsourced to commercial translation agencies
- Outsourcing models are changing:
 - Originally: pool of parties
 - then: single party
 - Currently: transitioning to a cascade model
 - Translations go to first preferred party
 - Overflow goes to second party, etc.

- Not just translation, also transcreation
- i.e. the new creation of a text in a different language
 - with contents similar to the source(s)
 - but no direct translation of these sources
 - (and the sources need not be text in all cases).
 - transcreation generally also involves localisation.

Cf. presentation by Michel Verhagen

- Document types:
 - speeches, speaking notes, memos, diplomatic cables, sometimes proposals for tender, especially for embassies (languages: NL, EN, FR, ES).
- Languages:
 - In the ministries: NL, EN, FR, ES
 - In the executing bodies (UWV, SVB): also other European languages and immigrant languages

Translation: Language pairs



- Total estimate: 350 million words per year
- Ministry of Foreign Affairs translates approx. 7 million words per year in house
- A standard contract (Foreigners chain, Ministry of Justice and Security)
 - clause to transfer TMs by default,
 - Often renegotiated to obtain a better price
- A standard contract (Ministry of Foreign Affairs):
 - does not mention translation memories at all.

- Some organisations use tools, some don't
 - [SDL Trados Studio](#)
 - MultiTerm 2019
 - Dragon Naturally Speaking 14
 - TextSTAT,
 - TermTreffer
 - WordBee or MultiTrans (SDL)
- [memoQ](#) is gaining in popularity
- Ministry of Foreign Affairs:
 - Used MT@EC and (currently) CEF eTranslation
 - But [DeepL](#) is considered better.
 - mainly for translation into French

- There is no data exchange at the national level
- There is no coordination at the national level
- Interinstitutionele Termraad ('Interinstitutional Term Council') for terminology
 - Terminological coordination and harmonisation between the EU translation services and the Dutch, Belgian and Flemish administrations
 - Representatives from:
 - European Court of Justice, European Commission, Council of the European Union, NL-TERM, European Court of Auditors, Ministry of Foreign Affairs (NL), Centre for Dutch-language terminology of the Dutch Language Union, General Secretariat of the Dutch Language Union, European Parliament, Translation Services Governance Department Services for General Government Policy (Flemish Government)
 - But it is currently not so active

- Organisations in NL:
 - [PIANOo](#) (Min Econ)
 - [English part](#)
 - [Tendered](#)
 - [RINIS](#) (Many Min)
 - [English part](#)
 - [ICTU](#) (Min Int)
 - [eDelivery](#)
 - [Logius](#) (Min Int)

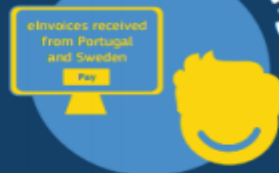
CEF-Building Blocks in NL:

- [eHealth](#) (GGZ)
- [eID](#) (Digibeter)
- [eInvoicing](#) (Min Int)
- [eSignature](#) (NL Enterprise Agency)
- [eDelivery](#) (ICTU)

eDelivery



eInvoicing



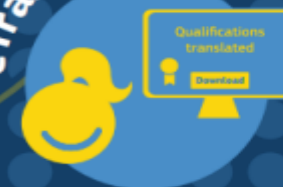
eID



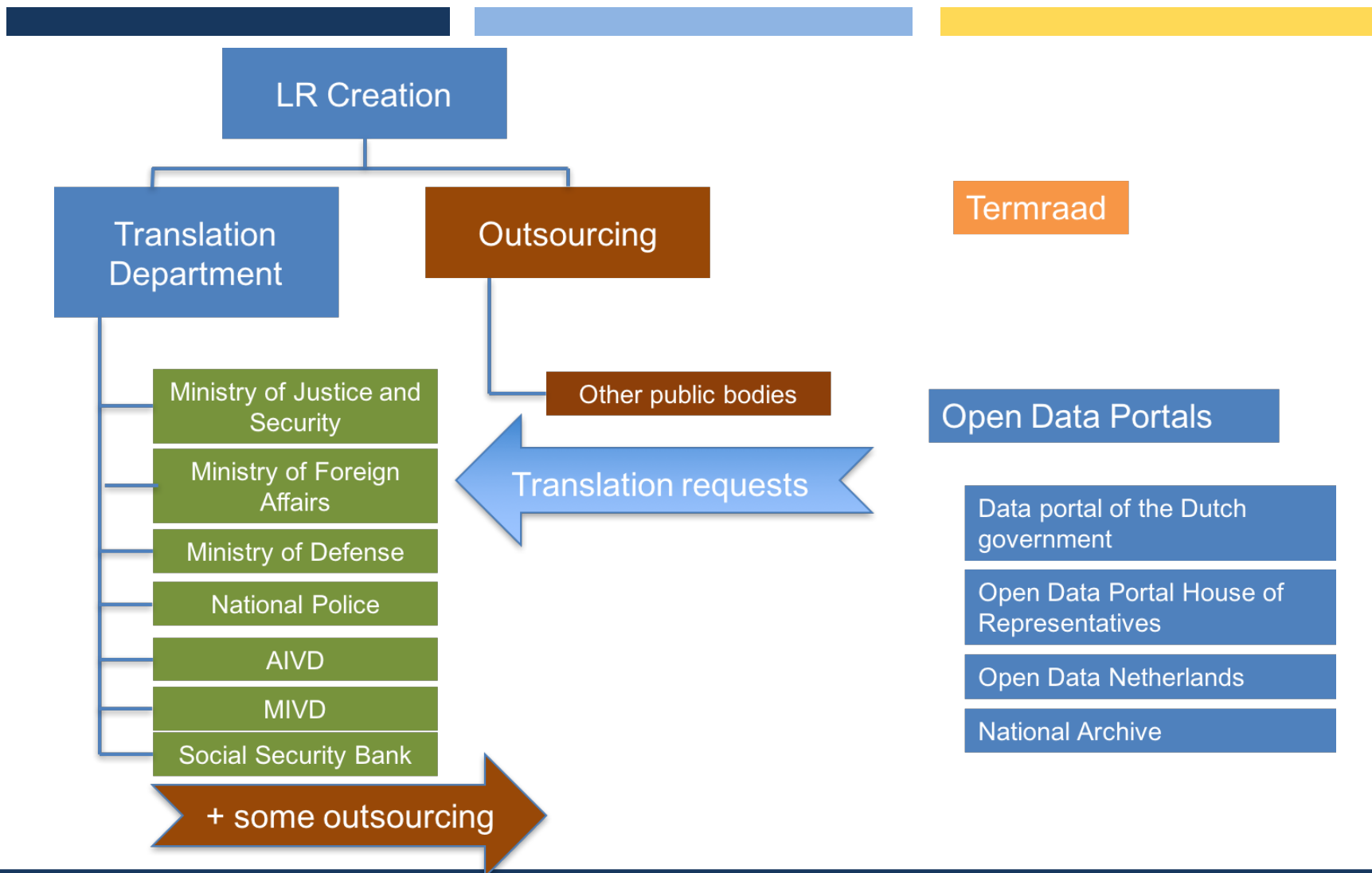
eSignature



eTranslation



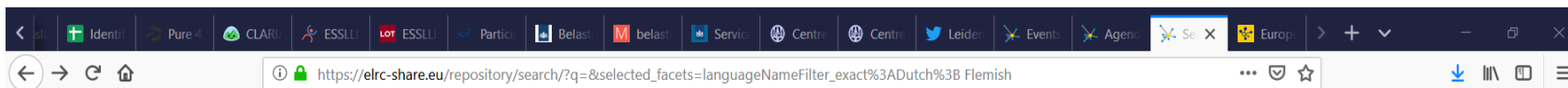
Infrastructure of language data on national level



- The Dutch Language
- Country Report
- **Identified Data**

- Datasets Min of Justice & Security: Procurement documents Translating / Interpreting
 - Monolingual
 - 2 batches
- Data set Social Security Bank (SVB)
 - Parallel Website texts in multiple languages
 - Sentence templates in multiple languages
- Data set Parliament: Texts of the Dutch Parliament
 - Monolingual
- Data set Dutch Language Union:
 - Parallel Corpus: nld, fra, eng
 - Multilingual subtitle data 2BDutch
 - Some monolingual corpora (SoNaR, DAESO)
- Data set Netherlands Authority for Consumers and Markets
 - Parallel texts Dutch - English
- Dutch texts and parallel corpora collected in Belgium
- Dutch texts of Hungarian websites

- Datasets Min of Justice & Security: Procurement documents
Translating / Interpreting
 - Monolingual
 - 2 batches
- Data set Dutch Language Union:
 - Parallel Corpus: nld, fra, eng
 - Multilingual subtitle data 2BDutch
 - Some monolingual corpora (SoNaR, DAESO)
- Data set Netherlands Authority for Consumers and Markets
 - Parallel texts Dutch - English
- Dutch texts and parallel corpora collected in Belgium
- Dutch texts of Hungarian websites



Search

Filter by:

- Language
 - Dutch; Flemish (29)
 - French (24)
 - English (15)
 - German (12)
 - Polish (7)
 - more
- Resource Type
- Media Type
- Licence
- Conditions of Use
- Validated
- Linguality Type
- Multilinguality Type
- Data Format

29 Language Resources (Page 1 of 2)

« Previous | Next »

Order by: Resource Name A-Z

- 2015 Calls for Tenders for Translation** 3 153

Dutch; Flemish

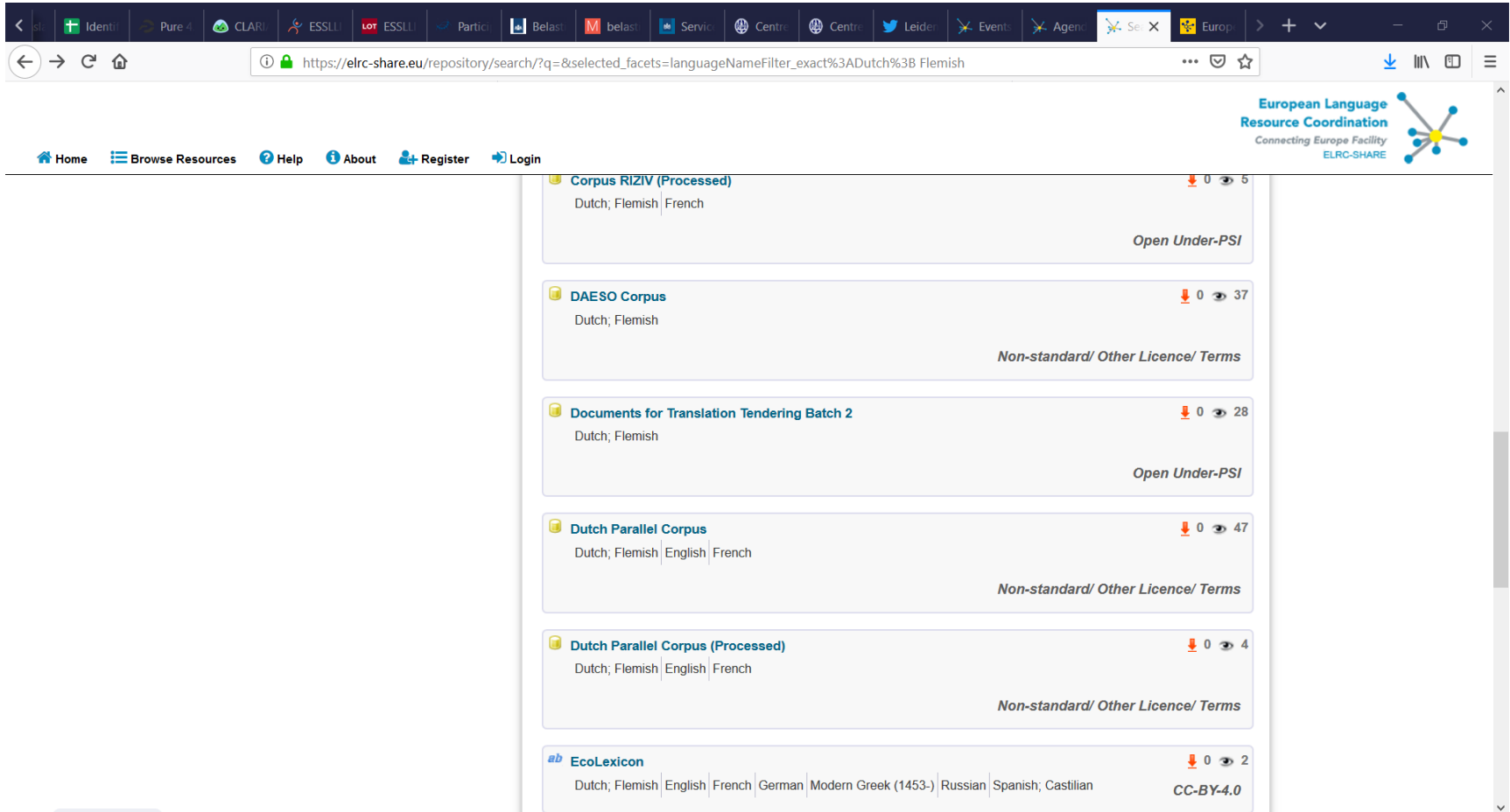
[Open Under-PSI](#)
- Belgian government bilingual parallel corpus** 0 80

Dutch; Flemish | French

[Open Under-PSI](#)
- Belgian government bilingual parallel corpus (Processed)** 0 9

Dutch; Flemish | French

[Open Under-PSI](#)



The screenshot shows a web browser window displaying search results for Dutch resources in the ELRC-SHARE repository. The search query is: `https://elrc-share.eu/repository/search/?q=&selected_facets=languageNameFilter_exact%3ADutch%3B Flemish`. The results are listed in a table-like format with the following entries:

Resource Name	Language(s)	Download Icon	View Icon	Count	License/Terms
Corpus RIZIV (Processed)	Dutch; Flemish French	📄	👁️	6	Open Under-PSI
DAESO Corpus	Dutch; Flemish	📄	👁️	37	Non-standard/ Other Licence/ Terms
Documents for Translation Tendering Batch 2	Dutch; Flemish	📄	👁️	28	Open Under-PSI
Dutch Parallel Corpus	Dutch; Flemish English French	📄	👁️	47	Non-standard/ Other Licence/ Terms
Dutch Parallel Corpus (Processed)	Dutch; Flemish English French	📄	👁️	4	Non-standard/ Other Licence/ Terms
EcoLexicon	Dutch; Flemish English French German Modern Greek (1453-) Russian Spanish; Castilian	📄	👁️	2	CC-BY-4.0

The page also features a navigation menu with links for Home, Browse Resources, Help, About, Register, and Login. The ELRC-SHARE logo is visible in the top right corner of the page content.

Results: new identified data (parallel and comparable)

Group	#Langs
AIVD	2
Business NL	2
Dutch Good Growth Fund	4
Dutch Tax Administration/business	3
Dutch Tax Administration/customs	2
Dutch Tax Administration/individuals	3
Forensic Institute	2
Government	4
Inspectie-OE	2
Medicine shortages and defects notification centre	2
Ministry of Defense	2
National Office for the Caribbean Netherlands	3
NederlandWereldwijd	6
Posted Workers	3
Relocatema	2
SVB	9
The Netherlands and You	12

Results: new identified data (parallel and comparable, by language)

Lang	Count
AR	1
Bahasa Indonesia	1
DE	6
EN	17
ES	4
FR	4
IT	1
JP	1
NL	17
Papiamentu	1
Papiamentu	2
PL	1
PT	3
RU	1
TR	2
ZH	1

- Other potential sources
 - Employee Insurance Agency (UWV)
 - University Websites
 - NWO (Science Foundation) websites and documents
 - European Sources (e.g. [EUR-LEX](#)), not considered here
 - [TenderNed](#)
 - digital announcements of calls for tenders.
 - fully connected to the European [Tenders Electronic Daily \(TED\)](#)
 - highly detailed legal and technical documents, summaries TenderNed.
 - in any language, usually Dutch, but English also occurs.
 - 60% structured metadata, 40% running natural language text

- How secure is the use of CEF eTranslation?
 - how secure is the connection? Can the security be tested / asserted somehow?
 - Are the data and translations deleted after sending them to the requester, or are they stored somewhere or even reused? Many data of some organisations that consider using eTranslation contain highly privacy-sensitive information.
- How can CEF eTranslate work better for me on the basis of my data? Will there be different versions? Or parameters to use different translation models? How does this work, especially when one and the same word or term must be translated differently for different organisations?
- If an organisation has relevant data to contribute, it will require effort from their side, sometimes significant effort because the data might be very distributed, and this forms an obstacle. Assistance by the ELRC team does not take this effort away. How can we address this problem?

Thank you for your attention!

Email: j.odijk@uu.nl; carole.tiberius@ivdnt.org

Website: <http://lr-coordination.eu/nl/l2netherlands>

