

Atelier ELRC en France

ELRC en France

François Yvon
LIMSI, CNRS



- Technologie des langues
 - Traduction automatique
 - Recherche d'information
 - Compréhension
 - Systèmes de dialogues
 - etc
 - Etat de l'art
 - Algorithmes d'apprentissage automatique (IA)
- + Ressources linguistiques et corpus annotés





Search

Filter by:

Language

- French (203)
- + English (170)
- + Spanish; Castilian (108)
- + German (98)
- + Catalan; Valencian (65)

more

- Resource Type
- Media Type
- Licence
- Conditions of Use
- Validated
- Linguality Type
- Multilinguality Type
- Data Format
- Domain
- Appropriateness For DSI
- Funding Project

203 Language Resources (Page 1 of 11)

« Previous | Next »

Order by: Resource Name A-Z

-  **ANR translation memory containing major publications, as well as several administrative documents and news** ↓ 25 👁 149

English | French *Open Under-PSI*
-  **ANR translation memory containing major publications, as well as several administrative documents and news (Processed)** ↓ 1 👁 17

English | French *Open Under-PSI*
-  **Apache Tika - a content analysis toolkit** ↓ 0 👁 8

Danish | Dutch; Flemish | English | Estonian | Finnish | French | German | Hungarian | Icelandic | Italian | Modern Greek (1453-) | Norwegian Bokmål | Polish | Portuguese | Spanish; Castilian | Swedish *Apache-2.0*
-  **Austrian Armed Forces Military Dictionaries** ↓ 17 👁 126

English | French | German | Hungarian | Italian *Open Under-PSI*

88 2 3

Belgian government bilingual parallel corpus

Attribution details: Ghent University

Aligned texts from the Belgian government in French and Dutch (aligned with SDL Trados Studio)

[← Back](#)

[Download](#)

Distribution

Availability: Available

Licences

Open Under-PSI

Used for resources that fall under the scope of PSI (Public Sector Information) regulations, and for which no further information is required or available. For more information on the EU legislation on the reuse of Public Sector Information, see here: <https://ec.europa.eu/digital-single-market/en/european-legislation-reuse-public-sector-information>.

Distribution Details

Attribution Details: Ghent University

Contact Person

[Ayla Rigouts Terryn](#) 

text 

Bilingual text corpus

Languages

French (fr)

Dutch; Flemish (nl)

Linguality

Linguality type: Bilingual

Text Format

TMX

TM format of the SDL alignment tool

Size

3,010 Translation Units

Character encoding

UTF-8

Resource Creation

Funding Project

Connecting Europe Facility - European Language Resource Coordination (CEF-ELRC - LANGUAGE RESOURCE COORDINATION - SMART 2014/1074 - 30-CE-0696785/00-64)

URL: <http://www.lr-coordi...>

Funding Type: Service Contract

Funder: European Commission

Funding Country: European Union (EU)

Project duration: 29/03/2015 - 16/04/2017

Metadata

Created: 08/12/2016

Last Updated: 08/12/2016

Metadata Language: English (en)

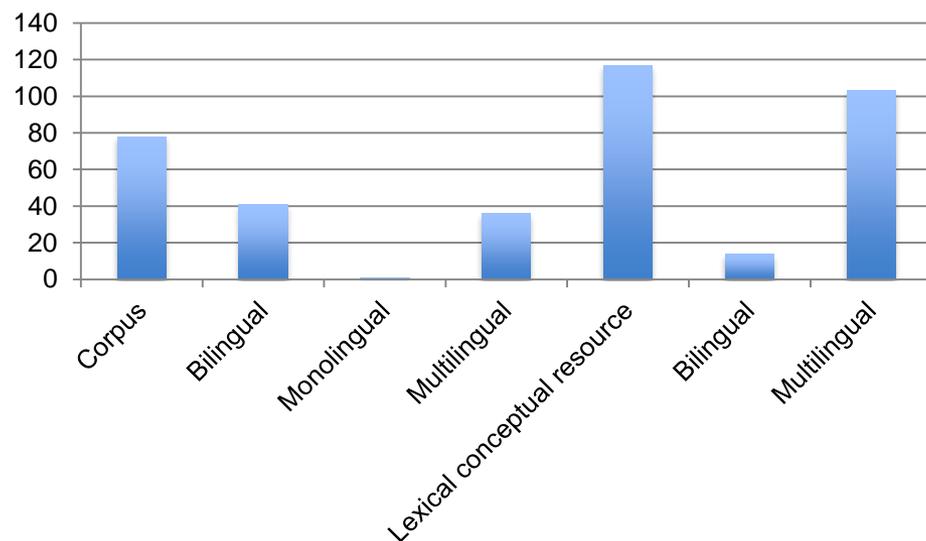
Version

Version: 1.0

Relations

Related Resource: [Belgian government bilingual parallel corpus \(Processed\)](#)

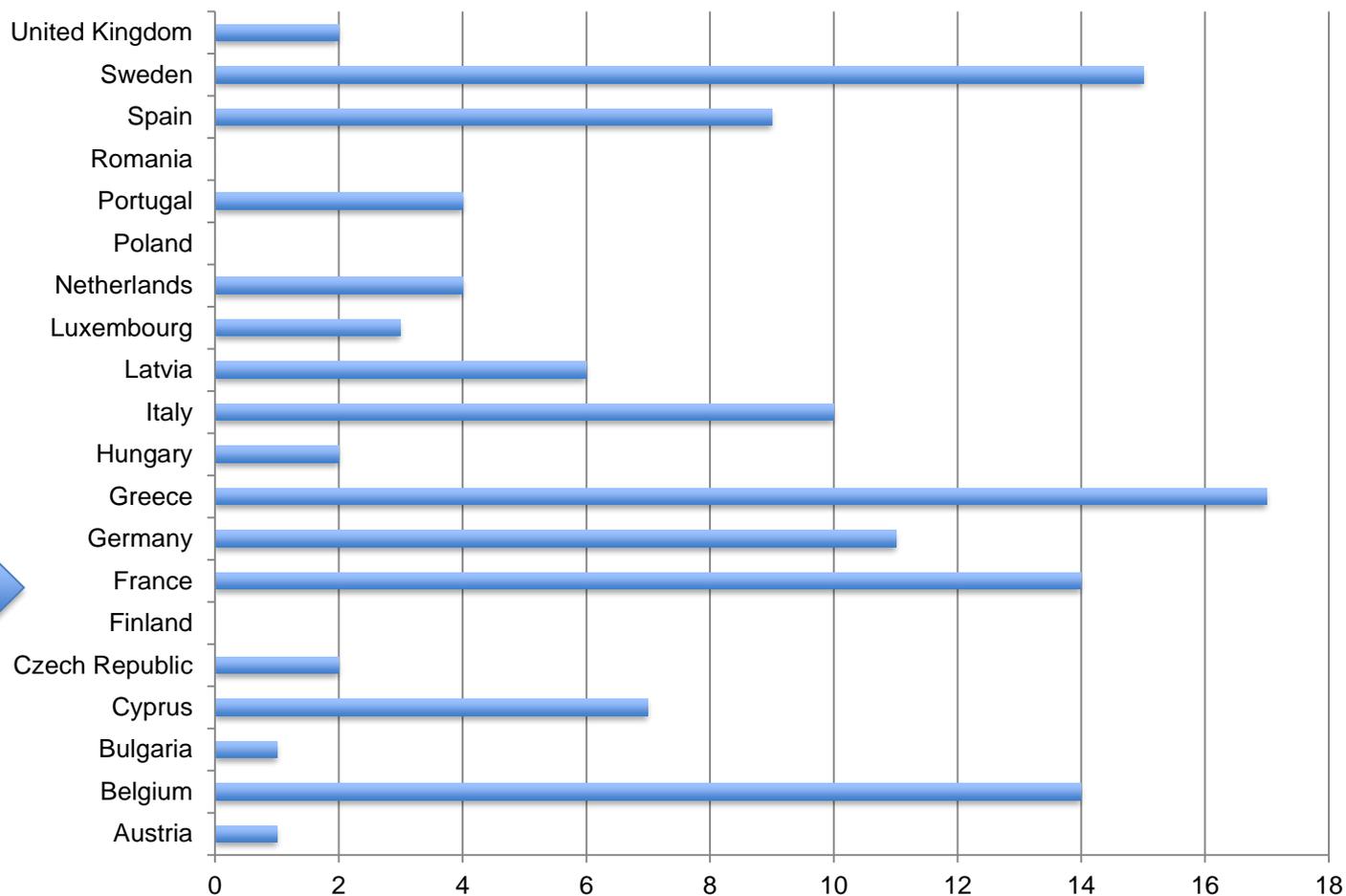
Relation Type: Has Aligned Version



ressources concernant le français

Corpus		
<i>Bilingues</i>		
Fichiers		61324
Unités de traduction		32864500
Mots		566405
<i>Monolingues</i>		
Unités de traduction		2295
<i>Multilingues</i>		
Unités de traduction		561011
Mots		427422
Lexiques, terminologies		
<i>Bilingues</i>		
Entrées		7093
<i>Multilingues</i>		
Entrées		277947
Fichiers		8
Termes		30633
Unités de traduction		837

L'origine des corpus par pays



Les principaux types de données collectées





Représentants nationaux (NAP)

- Représentant national service public : Thibault Grouas (DGLFLF)
- Représentant national technologique: François Yvon (CNRS)

High-level representatives involved

- Add name or position (institution)

Principaux fournisseurs de données

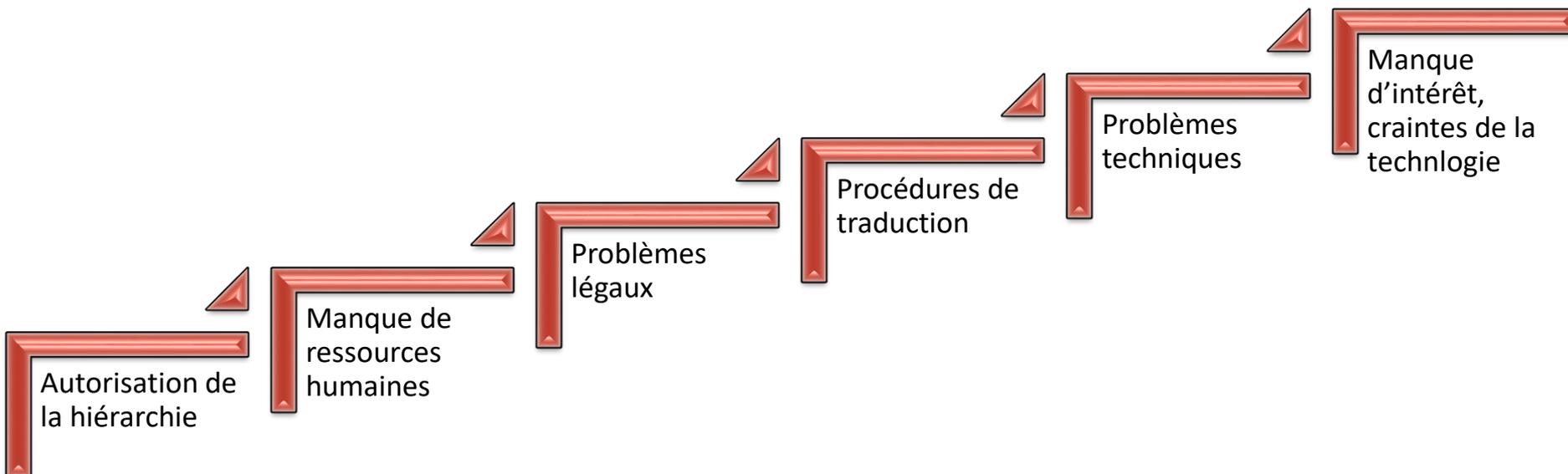
- Agence Nationale de la Recherche
- Ministère des finances / agence France Trésor
- Institutions étrangères (!)

Autres sources de données identifiées

- Participants à la réunion d'aujourd'hui



Principaux obstacles



Principales réalisations

- Quelques contributions très importantes
- De nombreux contributeurs de toute l'Europe
- Résolution des problèmes d'accès



Ce qui n'a pas marché

Comment améliorer la coordination nationale pour partager les bonnes pratiques et faciliter le partage de ressources ?

Améliorer les procédures de dépôt

Comment faire évoluer les procédures de traduction pour faciliter la réutilisation?

Autres améliorations possibles ?



Highlights and best practice - what worked well?

Key role / central player(s) and their support provided?

Modifications / improvements in translation procedures to facilitate data sharing?

Available languages?

Potential domains / areas covered?

Merci pour votre attention!

Courriel: info@lr-coordination.eu

Site web: www.lr-coordination.eu

