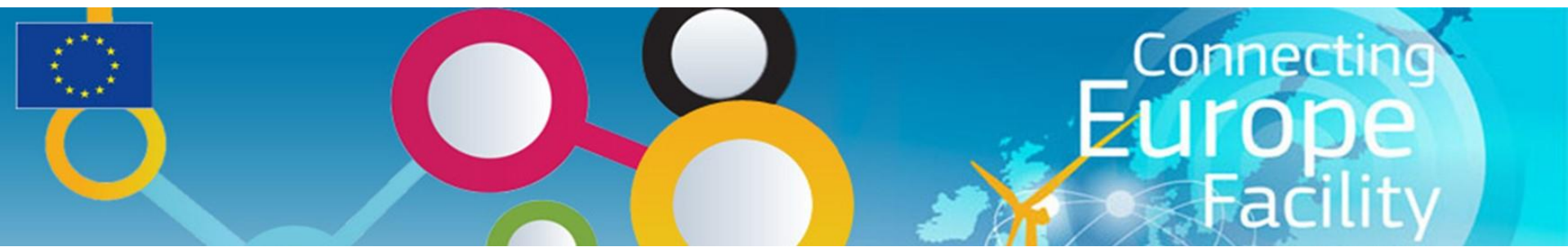


ELRC Workshop in Ireland

ELRC in Ireland

Micheál Ó Conaire

Department of Culture, Heritage and the Gaeltacht

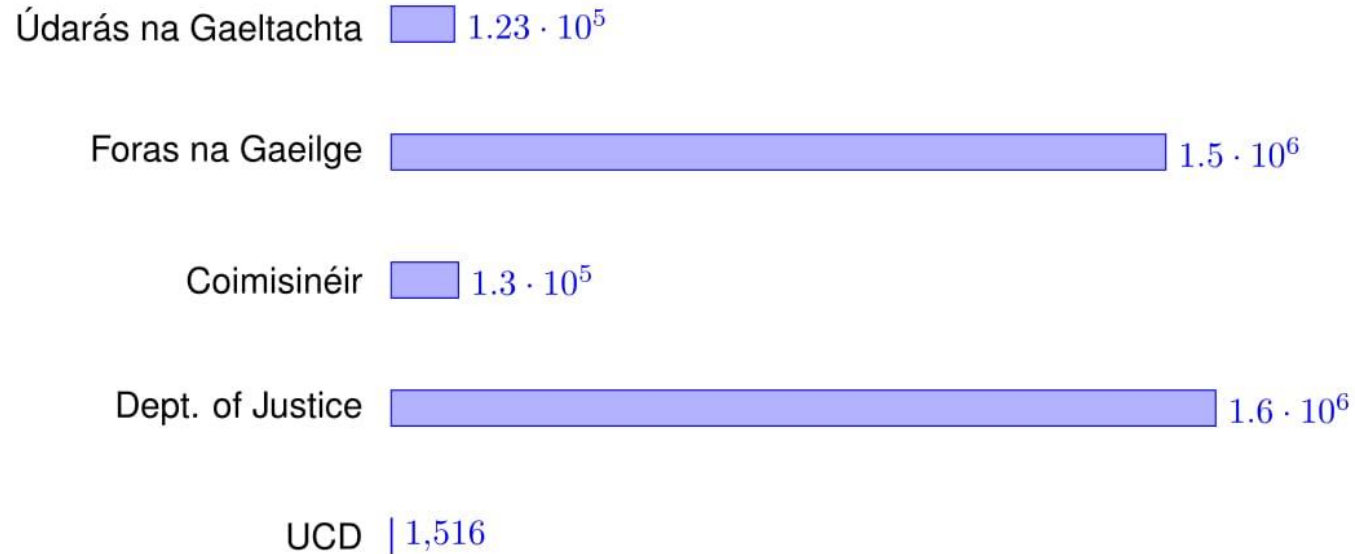




- The Gaois bilingual corpus of English-Irish legislation – languages: en/ga – format: txt – size: 1,596,569 TU
- The UCD Bórd na Gaeilge Corpus of bilingual PDFs and Word documents – languages: en/ga – format: PDF & Word – size: 1,516 TU
- The Coimisineir Teanga Bilingual Corpus of Reports and Press Releases – languages: en/ga – format: Word – size: 123,552 TU
- The Coimisineir Teanga Bilingual Web Corpus – languages: en/ga – format: TMX – size: 6,808 TU
- The Foclóir New English-Irish Dictionary – languages: en/ga – format: xml – size: xx 1,500,000
- The Udarás na Gaeltachta Corpus of bilingual PDFs and Word documents – languages: en/ga – format: PDF & Word – size: 122,547 TU



Number of Translation Units for Irish Resources





National Anchor Points

- Public Services NAP (institution)
- Technological NAP (institution)

High-level representatives involved

- Add name or position (institution)

Data contributors so far

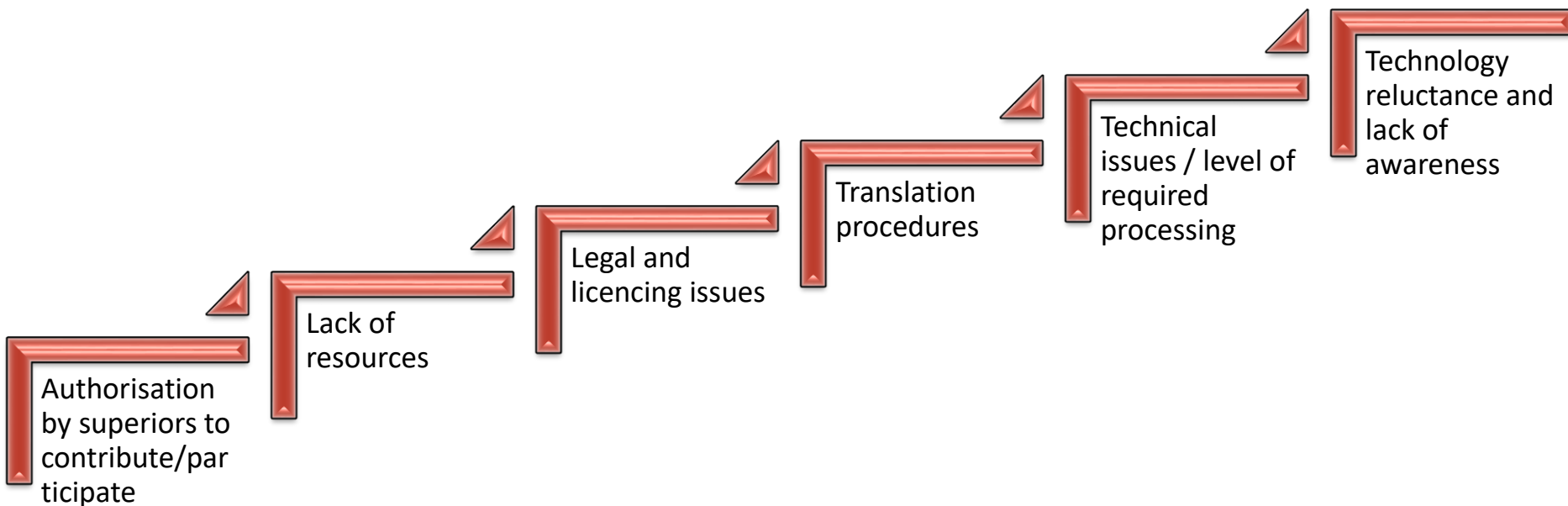
- UCD, Foras na Gaeilge, Dept. Of Justice, Údarás na Gaeltachta, An Coimisinéir Teanga

Potential data contributors / others

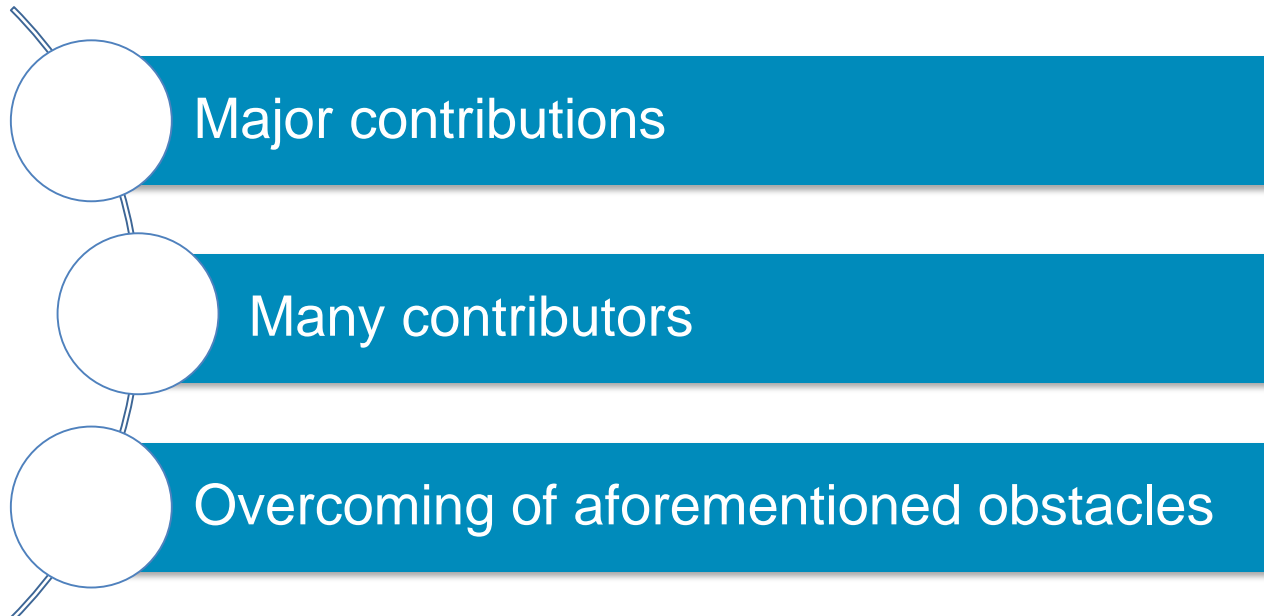
- Foras na Gaeilge, TCD, Conradh na Gaeilge, NÓS, The Irish Times



Main obstacles



Major achievements





Lessons learned - What did not work?

Explanation of data-types and ELRC objectives/requirements can be laborious via e-mail – slows down process

Turnaround of staff/ new equipment meaning original data not saved

Organisations not totally invested in data retention



Highlights and best practice - what worked well?

Foras na Gaeilge - knowledgeable about their own data and proactive about licencing.
Friendly and eager to help

Both monolingual and bilingual data gathered across a variety of domains and data-types

Web-crawling to supplement datasets, e.g An Coimisinéir Teanga

Phonecalls or in-house visits instead of e-mail chains

Thank you for your attention!

Email: info@lr-coordination.eu

Website: www.lr-coordination.eu

