

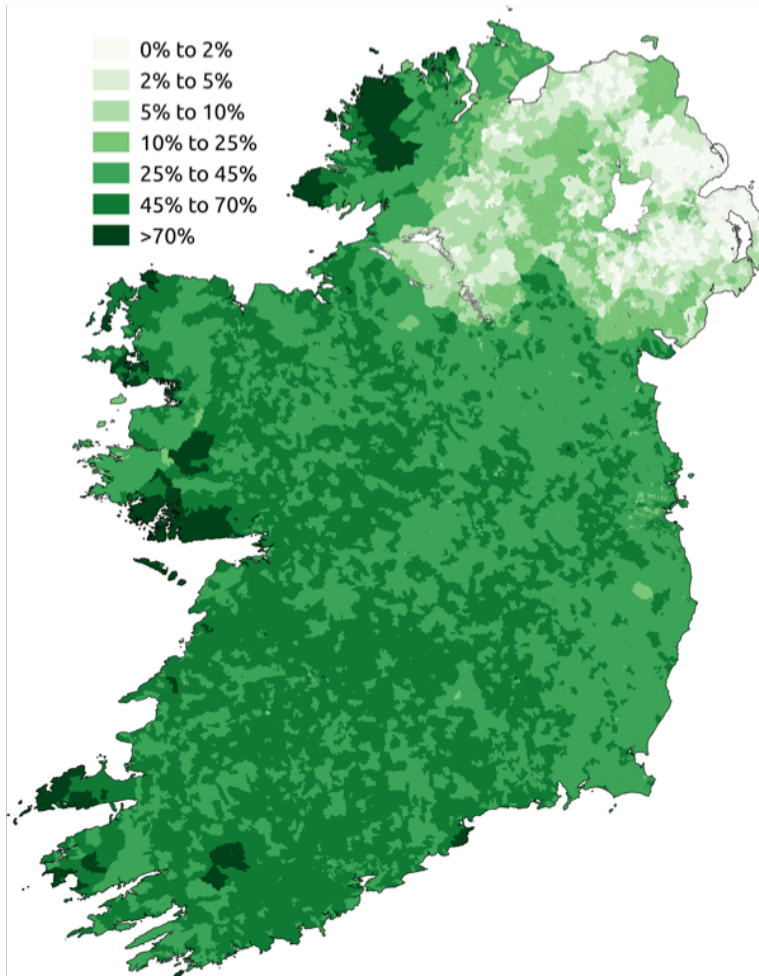


Engaging Content
Engaging People

Language Resource Collection in Ireland

Dr. Teresa Lynn
Research Fellow, ADAPT Centre
Dublin City University

- Background on Ireland's languages
- ELRC: National Anchor Point cooperation
- Future Translation Services plan
- ELRC – the Irish context

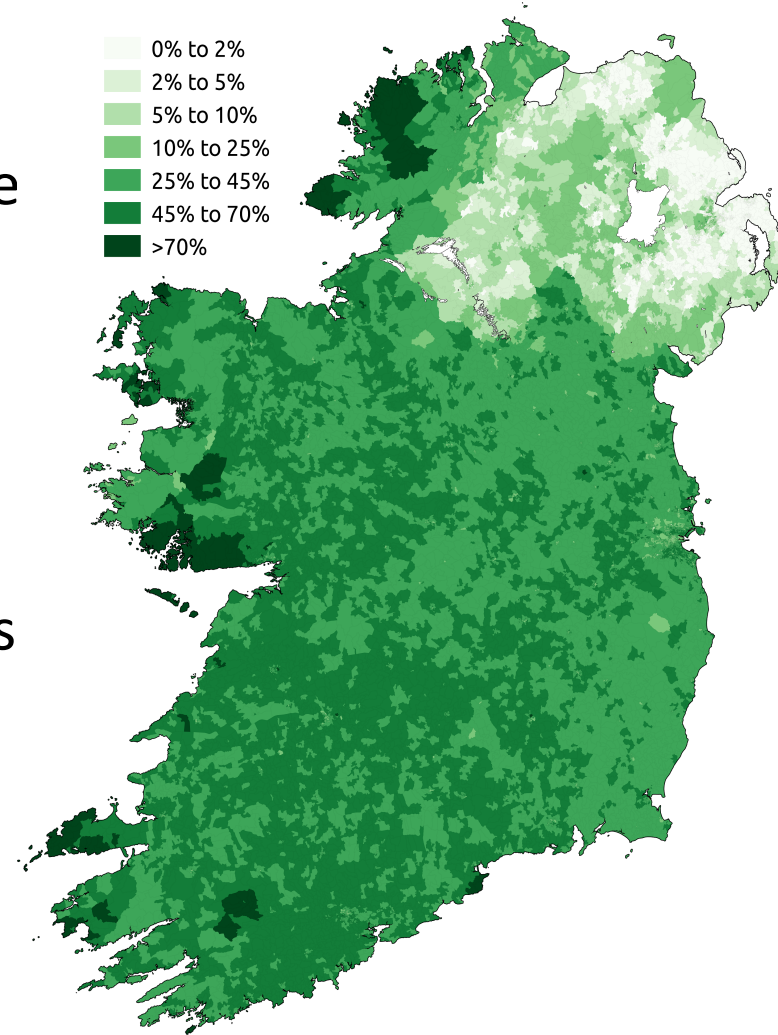


National Language
First Official Language
Official EU Language



Census: 2016
Population: 4,761,865
Daily usage: 73,803 people
Ability to speak: 1,761,420 people

- Once the most widely spoken language
- Historical factors led to decline in its use & increase in use of English
- Still spoken natively
 - Gaeltacht Regions
- “Urban revival” – Irish medium schools increasingly popular
- Increased online usage



- **1973** – Ireland joins EEC (EU): **Irish is a Treaty Language**, official procedural language of the European Court of Justice, correspond in written Irish with EU Institutions
- **1999** – British Irish Agreement Act: Recognised “the importance of respect understanding and tolerance in relation to linguistic diversity, the Irish language... **part of the cultural wealth of the island of Ireland**”
- **2003** – Official Languages Act: sets out rules regarding use of the Irish language by public bodies, **established the office of *An Coimisinéir Teanga* to monitor and enforce compliance** by public bodies
- **2007** – Official working language of the EU: Unanimous decision by EU foreign ministers (2005); **with a partial derogation** in place whereby only key legislation must be translated into Irish, i.e. Irish translations will only be carried out for documents covered by co-decision between the European Parliament and the Council of Ministers (**2021 – end of derogation**)
- **2012** – Gaeltacht Act: Updated language planning and policies, **gives statutory effect to a 20 year strategy for Irish**
- **2018** – Launch of National Digital Strategy for Irish

Public NAP

Technical NAP





An Roinn
Cultúir, Oidhreachta agus Gaeltachta

Department of
Culture, Heritage and the Gaeltacht

- **National Anchor Point:** Micheál Ó Conaire (Principal Officer)
- Government of Ireland Department/Ministry
- Oversees the protection and preservation of Ireland's heritage and cultural assets
- Promotes and protects heritage and culture
- Advances and supports the use of the Irish language
- Dedicated Minister of State with responsibility for Gaeltacht affairs (Irish language)

- **Initial meetings** between ADAPT and senior staff at DCHG (following META-NET White Paper Series)



- DCHG recognised **potential** for investing in language technology, specifically Machine Translation

- Potential **cost-savings** apparent



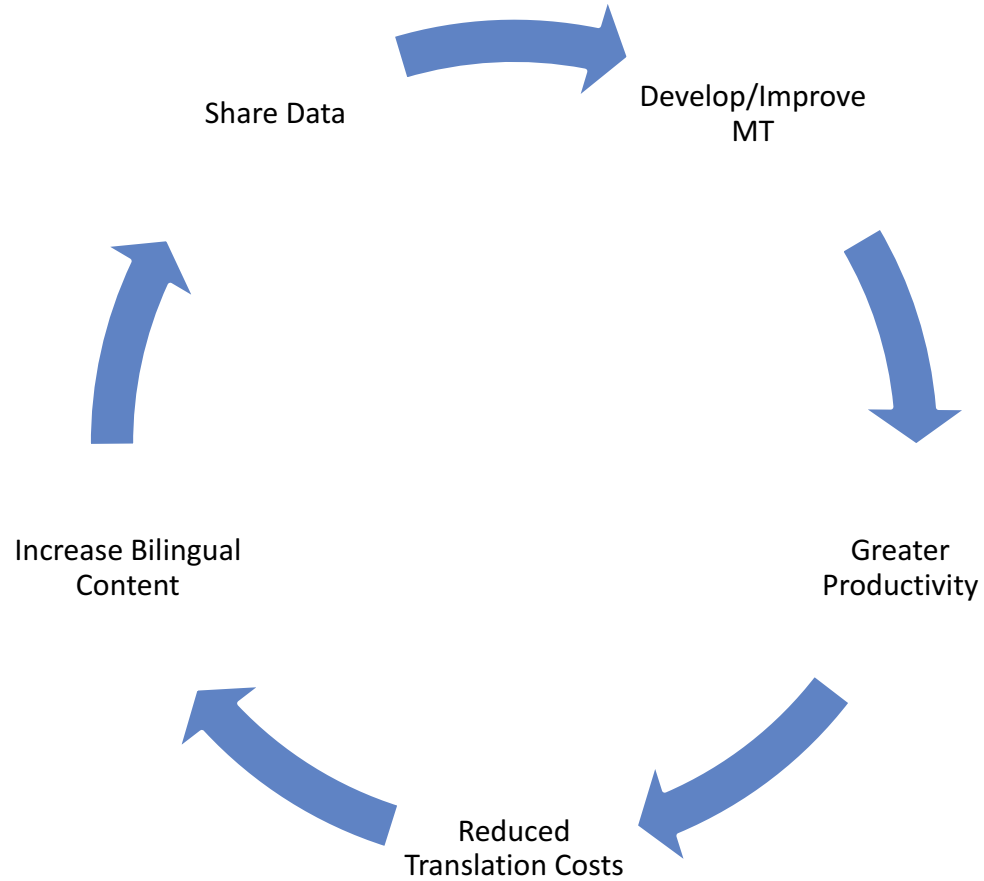
- National Digital Strategy (2018)
 - Similar to Digital Language Survival Kit



Public Service MT in Action

The screenshot displays a professional translation environment. The top toolbar is densely packed with icons for file management, editing, and translation-specific functions. The central workspace is divided into several panes. The main editor pane shows a document with a highlighted segment of text. To the right, a 'Term Recognition' pane lists terms found in the document. The bottom status bar provides real-time feedback on the translation quality, showing a score of 53.29% for the current segment.





A Shared Translation Service



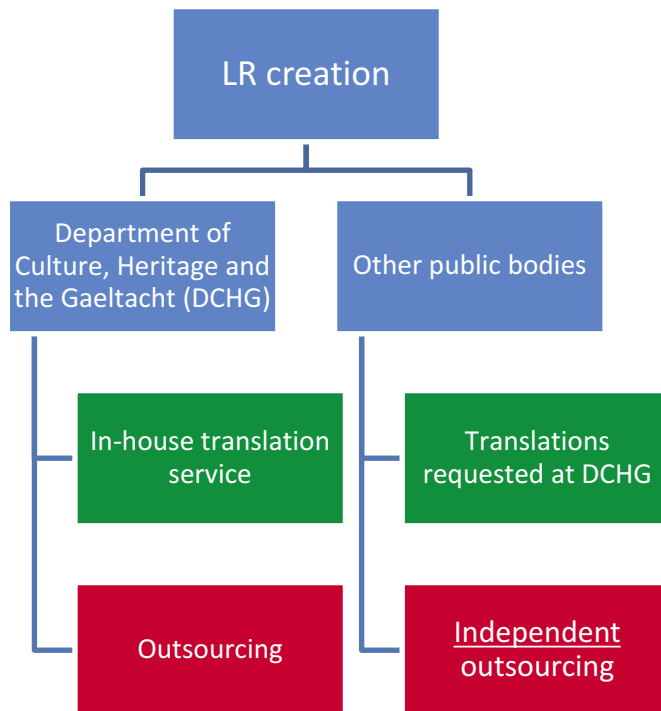
- Government Decision (June 2017) – Official Languages (Amendment) **Bill**
- Authorised initiatives including the establishment of a **Shared Translation Service**
- STS will operate within DCHG’s Irish Language Policy division and be available for **all government departments**
- Scale STS to **other languages** for other Departments
eg. Health, Justice, Social Protection, Children

*“... Central to the successful implementation of this plan and to the achievement of value for money and economies of scale is **the development of an effective machine translation system.** ...”*

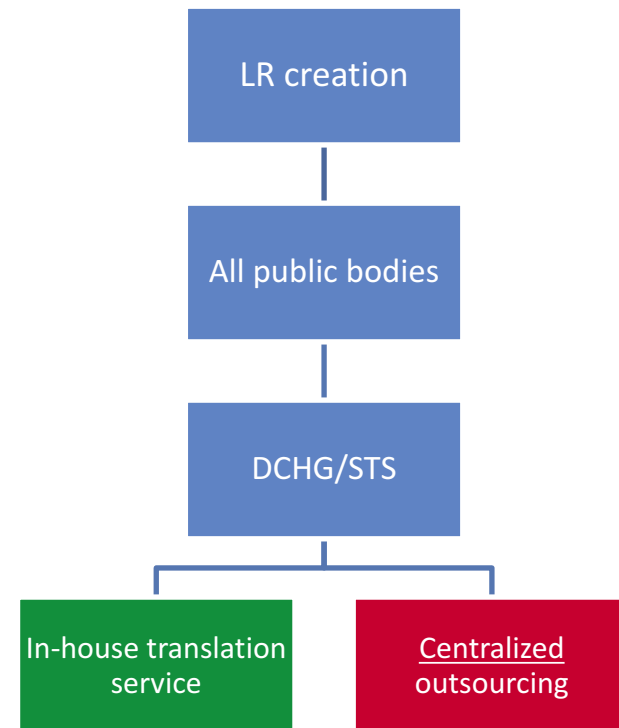
- DCHG report to Government on the STS

Infrastructure of language data on National level (Ireland)

CURRENT STATUS



GOAL



Advantages of Irish context

- Small country – less public bodies/ organisations to reach
- Bilingual Society
- Dedicated Government Department for Irish language affairs
- Dedicated Terminology Committee & Database
- Easier to make personal connections
- Department of Gaeltacht holds influence/sway
- Positive disposition towards language support

Advantages within the Irish context

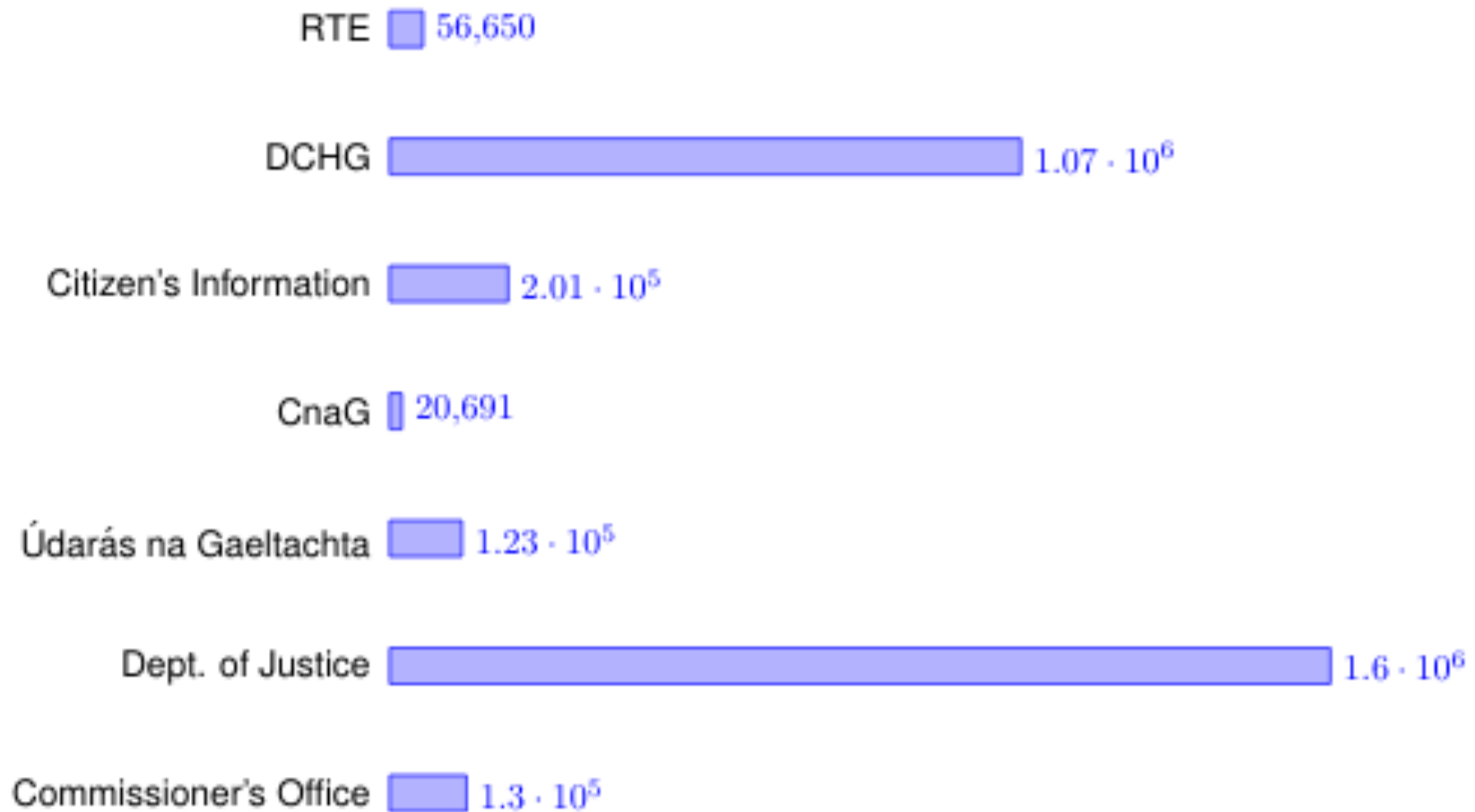
- Small country – less public bodies/ organisations to reach
- Bilingual Society
- Dedicated Government Department for Irish language affairs
- Dedicated Terminology Committee & Database
- Easier to make personal connections
- Department of Gaeltacht holds influence/sway
- Positive disposition towards language support

Shared challenges within the Irish context

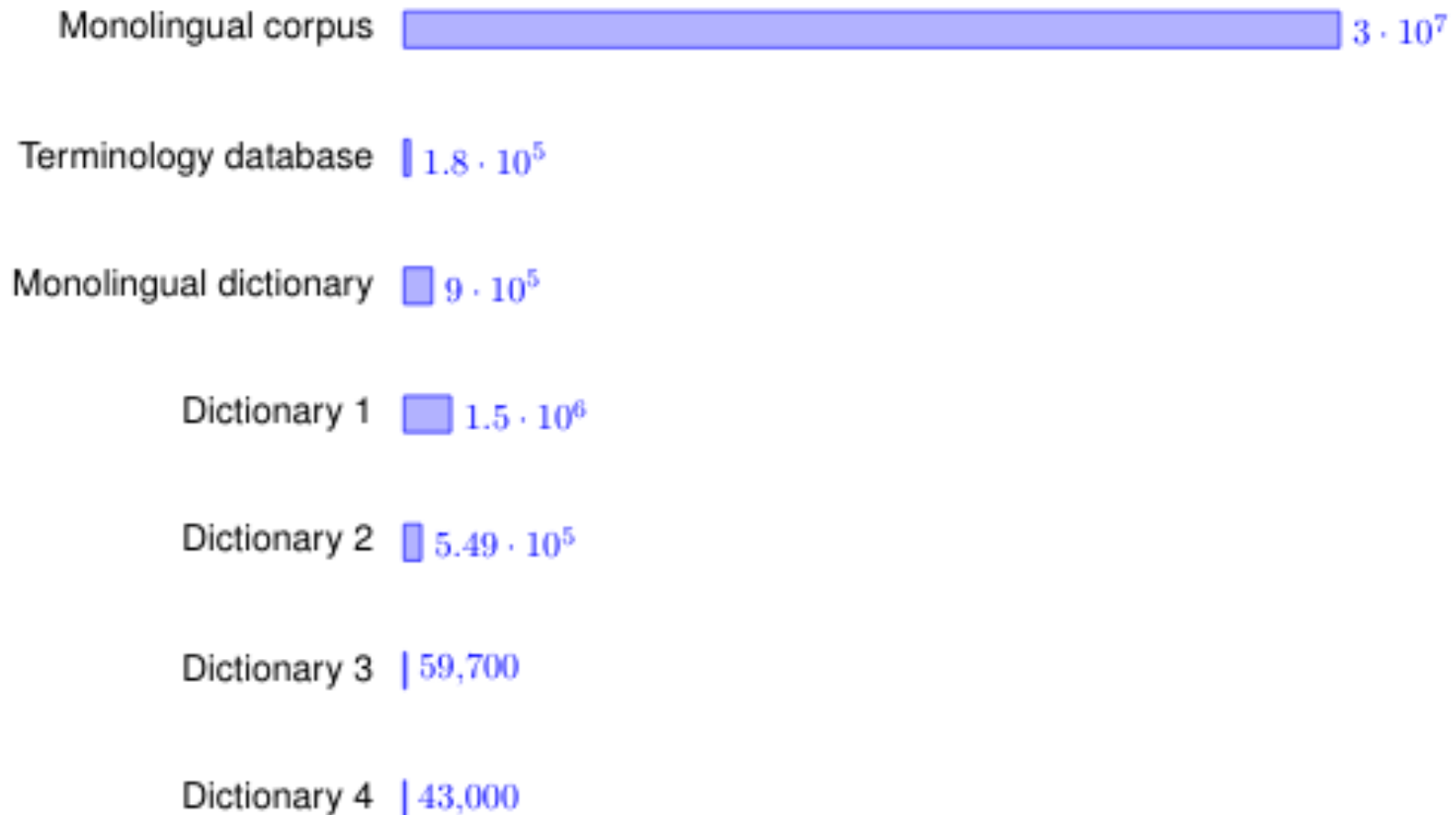
- Lack of knowledge about value of language data
- Lack of language technology use
- Inefficient data management practices
- Frequent changes in Ministerial structures
- Difficult to ascertain “who is doing what?”
- Dealing with apathy within public service/ demarcation issues




Number of Irish Words/Entries



Foras na Gaeilge: Number of Irish Words/Entries




Foras na Gaeilge: Number of Irish Words/Entries


Terminology database  $1.8 \cdot 10^5$

Monolingual dictionary  $9 \cdot 10^5$

Dictionary 1  $1.5 \cdot 10^6$

Dictionary 2  $5.49 \cdot 10^5$

Dictionary 3  59,700

Dictionary 4  43,000

- **Dept of Culture, Heritage and the Gaeltacht**
Through developing MT engine – showed the *value* of data sharing
- **Dept of Culture, Heritage and the Gaeltacht**
Dept of Justice
Dept of Communications, Climate Action and Environment
Through providing advice on how to *negotiate* translation contracts
(ownership of TMX files)

“Where computer aided translation tooling is used to produce translations then, in addition to the finished target language translation, any and all translation memory data, or similar, derived from the translation process should be returned to the department contracting for translation services as part of the final end product delivery.”





- Ireland, Spain, France and Portugal
- Collect, Prepare and Share Language Resources
- File alignments, TMX creation
- National Relay Stations



- Dr. Teresa Lynn
Research Fellow
ADAPT Centre, Dublin City University
teresa.lynn@adaptcentre.ie
- **ELRC NAPs Ireland**
 - Technical : Prof. Andy Way,
ADAPT Centre, Dublin City University
andy.way@adaptcentre.ie
 - Public Service: Micheál Ó Conaire,
Department of Culture, Heritage and the Gaeltacht
micheal.oconaire@chg.gov.ie