The European Language Resource Coordination (ELRC)

Stelios Piperidis
Institute for Language and Speech Processing / Athena R.C. ELRC
The ELRC context
The ELRC context
Heads & organisations behind the ELRC

The ELRC Consortium:
Heads & organisations behind the ELRC

30 ELRC Technological National Anchor Points

30 ELRC Public Services National Anchor Points
The ELRC anchor points in Romania

Laura Mihăilescu
Translation Coordination Unit, European Institute of Romania

Dan Tufiș
Research Institute for Artificial Intelligence, Romanian Academy of Sciences
What does ELRC do?

- **Collect**
  - Language resources

- **Identify**
  - Needs of public services

- **Engage**
  - With the public sector in the identification of LR

- **Help**
  - With any technical or legal issues

- **Act**
  - Observatory for language resources across Europe
Why ELRC...?

• To facilitate multilingual communication and exchange of information in key public service scenarios
• How? With the right language resources!
• Examples of public services scenarios to which eTranslation may be applied:
  – Consumer Rights (ODR),
  – Social Security (EESSI),
  – Justice (eJustice),
  – Commerce (BRIS),
  – Open Data (EDP),
  – ...

The European Parliament voted in favor of the resolution on "Language equality in the digital age" on 11. September
“18. Calls on the Commission and the Member States to develop strategies and policy action to facilitate multilingualism in the digital market; requests, in this context, that the Commission and the Member States define the minimum language resources that all European languages should possess, such as data sets, lexicons, speech records, translation memories, annotated corpora and encyclopaedic content, in order to prevent digital extinction;”
“21. Calls on the Commission to make as a priority of language technology those Member States which are small in size and have their own language, in order to pay heed to the linguistic challenges that they face;”
Committee on Industry, Research and Energy (11.9.18): stresses "the role of the EU, the Member States, universities and other public institutions in contributing to the preservation of their languages in the digital world and in developing databases and translation technologies for all EU languages, including languages that are less widely spoken; calls for coordination between research and industry with a common objective of enhancing the digital possibilities for language translation and with open access to the data required for technological advancement".
Why ELRC...?

• Cracking the Language Barrier Federation (2016): “We recommend that Europe actively makes an effort to compete in the global landscape for research and development in language technology since we cannot expect third parties from other continents to solve our translation and knowledge management problems in a way that suits our specific communicative, societal and cultural needs. Language Technology made for Europe in Europe is the key. It will contribute to future European crossborder and cross-language communication, economic growth and social stability.”
What does ELRC do?

**Collect**
- Language resources

**Identify**
- Needs of public services

**Engage**
- With the public sector in the identification of LR

**Help**
- With any technical or legal issues

**Act**
- Observatory for language resources across Europe
How to make it work...?

With the right language resources!

- Add in-domain text
- In-domain data
- Good translation
How to make it work...?

In-domain data → Out-of-domain data

→ Your language resources make a difference!
What has been achieved so far…?

The ELRC network

More than 400 identified data holders

60 National Anchor Points

ELRC consortium

Today

April 2016

April 2015
What has been achieved so far...?

ELRC Data in April 2017

A total of 225 LR provided by ELRC

- Open data / cc licenses: 78%
- Non-standard/other licenses: 17%
- Under review: 5%

[Pie chart showing the distribution of license types]
What has been achieved so far...?

ELRC Data in October 2018

More than 500 LRs contributed through ELRC

- 73% cc / open licences
- 15% non-standard licences
- 12% under review
What has been achieved so far...?

ELRC Data by language

Development of LR collection by language

Apr 17  Oct 18
What has been achieved so far...?

ELRC Data by language

Development of LR collection by language

- BG, CS, DA, DE, EL, EN, ES, ET, FI, FR, GA, HR, HU, IS, IT, LT, LV, MT, NL, NO, PL, PT, RO, SK, SL, SV

- Apr 17 and Oct 18
• **The ELRC-SHARE Repository**
  – Access to, sharing and contribution of language resources
  – Access to tools and services catalogue (upcoming)
  – Visit [https://elrc-share.eu](https://elrc-share.eu)
• **ELRC Technical and Legal Helpdesk**
  
  – Continuous support for potential data contributors (phone, email, ...)
  
  – Accessible online via [http://www.lr-coordination.eu/helpdesk](http://www.lr-coordination.eu/helpdesk)
• **ELRC On-site assistance**
  – Submit request online at [http://www.lr-coordination.eu/services](http://www.lr-coordination.eu/services)
Thank you for your attention!

Email: info@lr-coordination.eu
Website: www.lr-coordination.eu