

ELDA at the crossroads: bridging language resources and responsible AI

Gabriele Chignoli
LREC 2026 – Industry day
14/5/2026

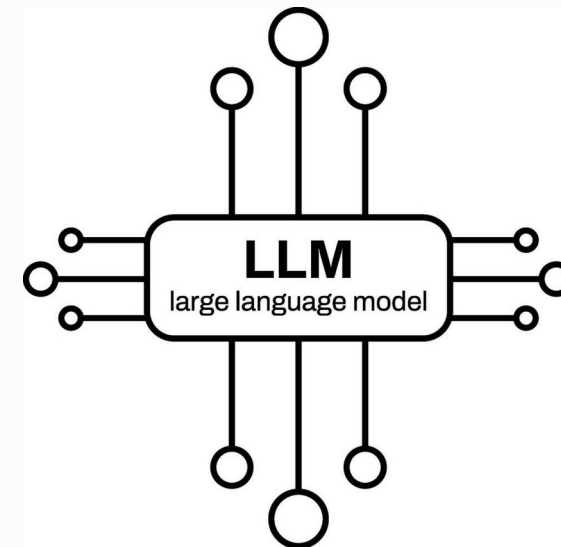




Introduction

The AI Revolution: A Crisis of Trust

- LLMs transformed AI overnight
- But where did they get their answers?
- Data is often opaque, biased, non-compliant





Introduction

1995 → 2026

- 30 years building high-quality resources
- Bridging linguistic excellence and AI revolution
- Next step: better AI, not just more AI



What is ELDA

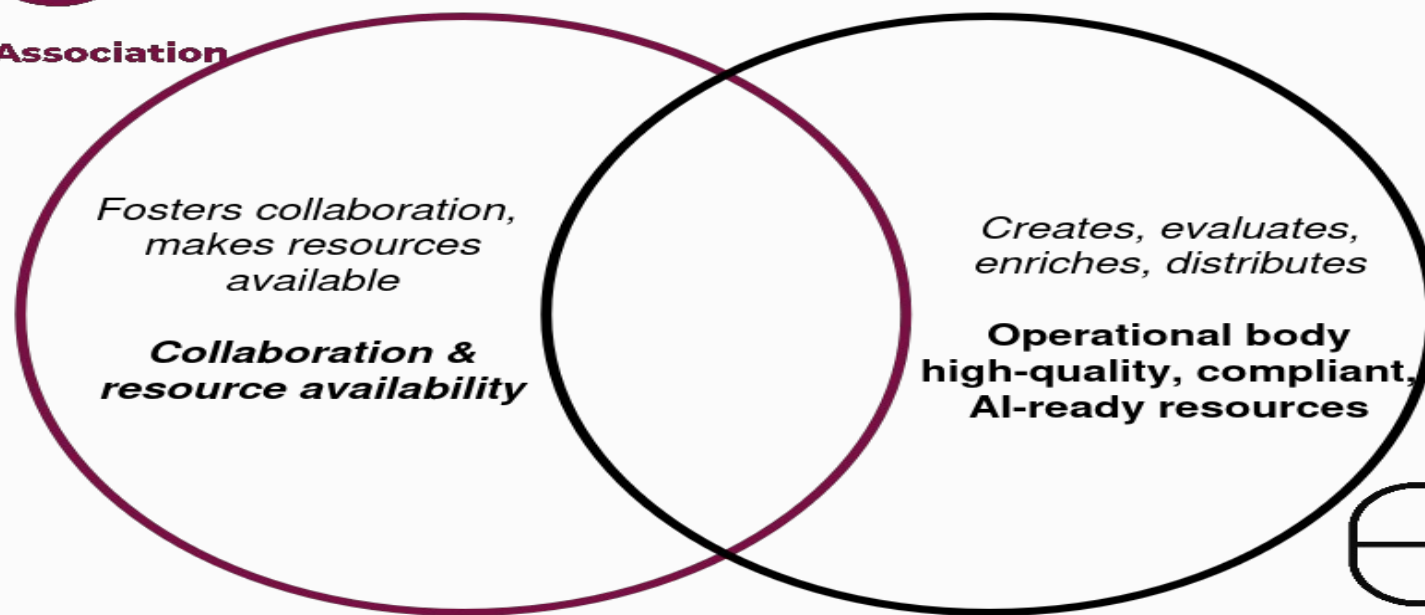


What is ELDA

Our role and relationship with ELRA

elra

Language Resources Association





What is ELDA

Full lifecycle of language resources

- From legacy to future-proof
- Raw data & repurposing
- LLM fine-tuning, evaluation, domain adaptation





What we do



What we do

How we curate, evaluate, and shape the future of linguistic AI





What we do

Trsproc

- Open-source tool
- Simplify and scale annotation based on the TRS format
- Ensuring future-proof of legacy resources

The screenshot shows the PyPI page for the `trsproc` package, version 2.1.0. The page has a blue header with the package name and version. Below the header, there is a navigation sidebar on the left with links to 'Project description', 'Release history', and 'Download files'. The main content area on the right is titled 'Project description' and contains a 'README' section. The README text describes `trsproc` as a Python module for processing TRS files from Transcriber and provides installation instructions. A green checkmark icon and the text 'Latest version' are visible in the top right corner of the header area. The release date 'Released: Apr 16, 2026' is also displayed.



Where we are headed



Where we are headed

Our vision for the coming years

- Trustworthiness
- Data Sovereignty
- Inclusive Innovation



To sum up

Collective responsibility to:

- Building and maintaining the linguistic data ecosystem
- Collaborate and share best practices