



**EUROPEAN
LANGUAGE
EQUALITY**

META  **NET**
META  **FORUM 2023**

New Features in the Digital Language Equality Dashboard and an Update of the Cross-language Comparison

Maria Giagkou, Stelios Piperidis, Athanasia Kolovou, Miltos Deligiannis, Leon Voukoutis, Penny Labropoulou
(Institute for Language and Speech Processing / RC Athena)

27-06-2023 META-FORUM 2023 – Digital Language Equality for a Multilingual Europe
<http://european-language-equality.eu>

Outline

- Recap: ELE 2021-22
- The ELG Catalogue as source of evidence
- The Digital Language Equality dashboard
- Cross-language comparison update





Recap: ELE 2021-22

ELE 2021-22

Some of the main outcomes and findings

- Evidence-based investigation of the level of technology support per language
- Aggregating (meta)data through multiple strategies, including manual population by ELE partners
- Use of the ELG Catalogue as database for further DLE computations and visualisations through the DLE dashboard (<https://live.european-language-grid.eu/catalogue/dashboard>)



ELE 2021-22

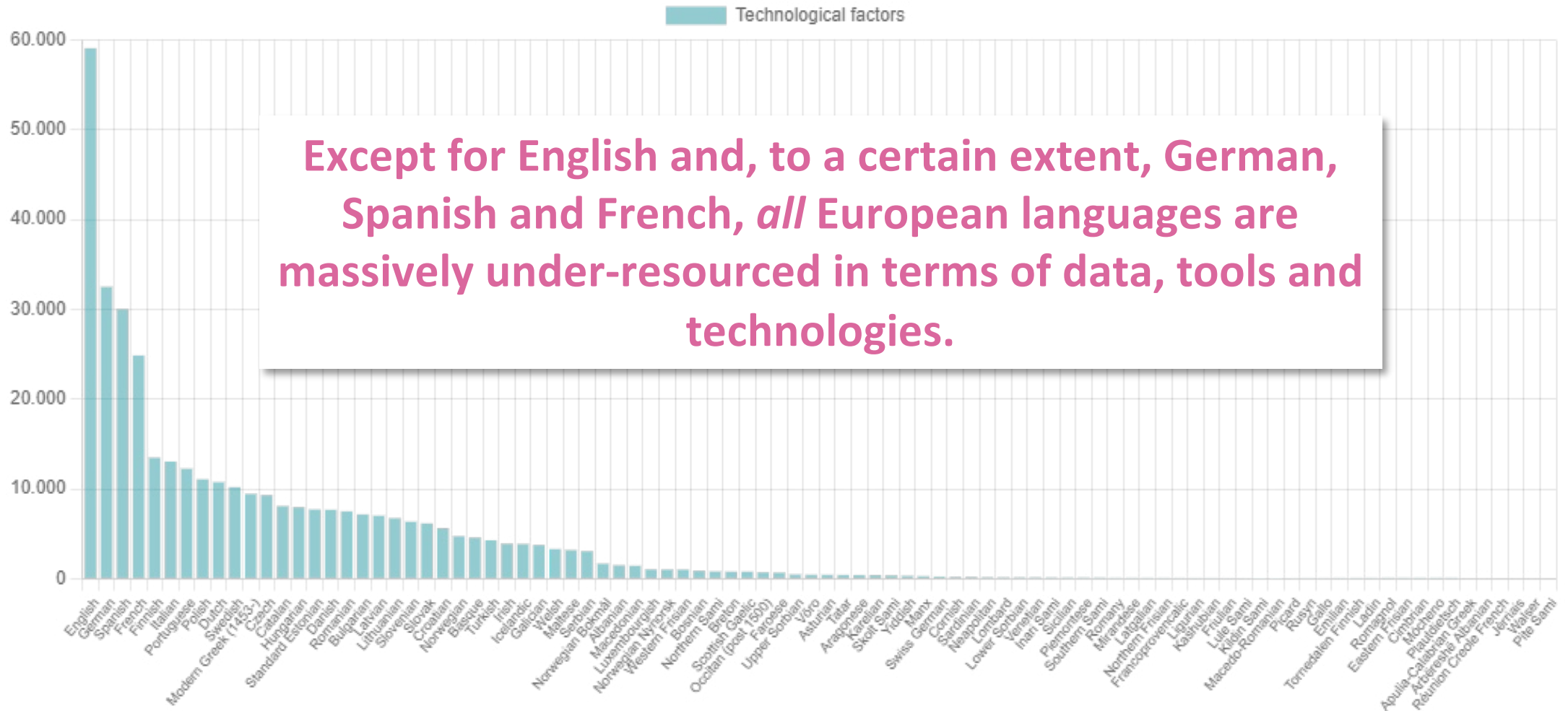
Some of the main outcomes and findings

- 35 language reports
- Cross-language comparison:
 - EU official languages better supported than other European languages
 - Best supported: English, followed by Spanish, German and French
 - Least supported among the EU official: Irish and Maltese
 - Translation technologies: many languages are at least moderately supported
 - Multimodal corpora and Models: many languages are weakly supported

| | | Tools and Services | | | | | | | Language Resources | | | | | |
|-------------------------|----------------|--------------------|-------------------|------------------------|-------------------------------|----------------------------|--------------------------|-----------------------------|--------------------|--------------------|------------------|--------|-------------------|---------|
| | | Text Processing | Speech Processing | Image/Video Processing | Information Extraction and IR | Human-Computer Interaction | Translation Technologies | Natural Language Generation | Text Corpora | Multimodal Corpora | Parallel Corpora | Models | Lexical Resources | Overall |
| EU official languages | Bulgarian | | | | | | | | | | | | | |
| | Croatian | | | | | | | | | | | | | |
| | Czech | | | | | | | | | | | | | |
| | Danish | | | | | | | | | | | | | |
| | Dutch | | | | | | | | | | | | | |
| | English | | | | | | | | | | | | | |
| | Estonian | | | | | | | | | | | | | |
| | Finnish | | | | | | | | | | | | | |
| | French | | | | | | | | | | | | | |
| | German | | | | | | | | | | | | | |
| | Greek | | | | | | | | | | | | | |
| | Hungarian | | | | | | | | | | | | | |
| | Irish | | | | | | | | | | | | | |
| | Italian | | | | | | | | | | | | | |
| | Latvian | | | | | | | | | | | | | |
| | Lithuanian | | | | | | | | | | | | | |
| | Maltese | | | | | | | | | | | | | |
| | Polish | | | | | | | | | | | | | |
| | Portuguese | | | | | | | | | | | | | |
| | Romanian | | | | | | | | | | | | | |
| Slovak | | | | | | | | | | | | | | |
| Slovenian | | | | | | | | | | | | | | |
| Spanish | | | | | | | | | | | | | | |
| Swedish | | | | | | | | | | | | | | |
| (Co-)official languages | National level | Albanian | | | | | | | | | | | | |
| | | Bosnian | | | | | | | | | | | | |
| | | Icelandic | | | | | | | | | | | | |
| | | Luxembourgish | | | | | | | | | | | | |
| | | Macedonian | | | | | | | | | | | | |
| | | Norwegian | | | | | | | | | | | | |
| | Regional level | Serbian | | | | | | | | | | | | |
| | | Basque | | | | | | | | | | | | |
| | | Catalan | | | | | | | | | | | | |
| | | Faroese | | | | | | | | | | | | |
| | | Frisian (Western) | | | | | | | | | | | | |
| | | Galician | | | | | | | | | | | | |
| | | Jerriais | | | | | | | | | | | | |
| | | Low German | | | | | | | | | | | | |
| | | Manx | | | | | | | | | | | | |
| | | Mirandese | | | | | | | | | | | | |
| | | Occitan | | | | | | | | | | | | |
| | | Sorbian (Upper) | | | | | | | | | | | | |
| | | Welsh | | | | | | | | | | | | |
| All other languages | | | | | | | | | | | | | | |

ELE 2021-22

Some of the main outcomes and findings

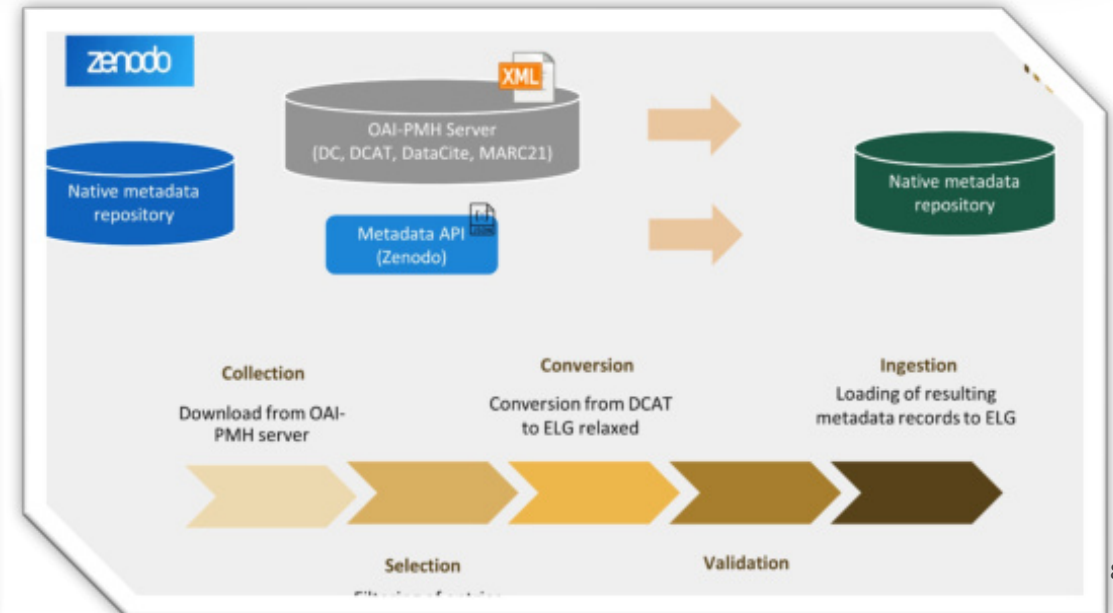
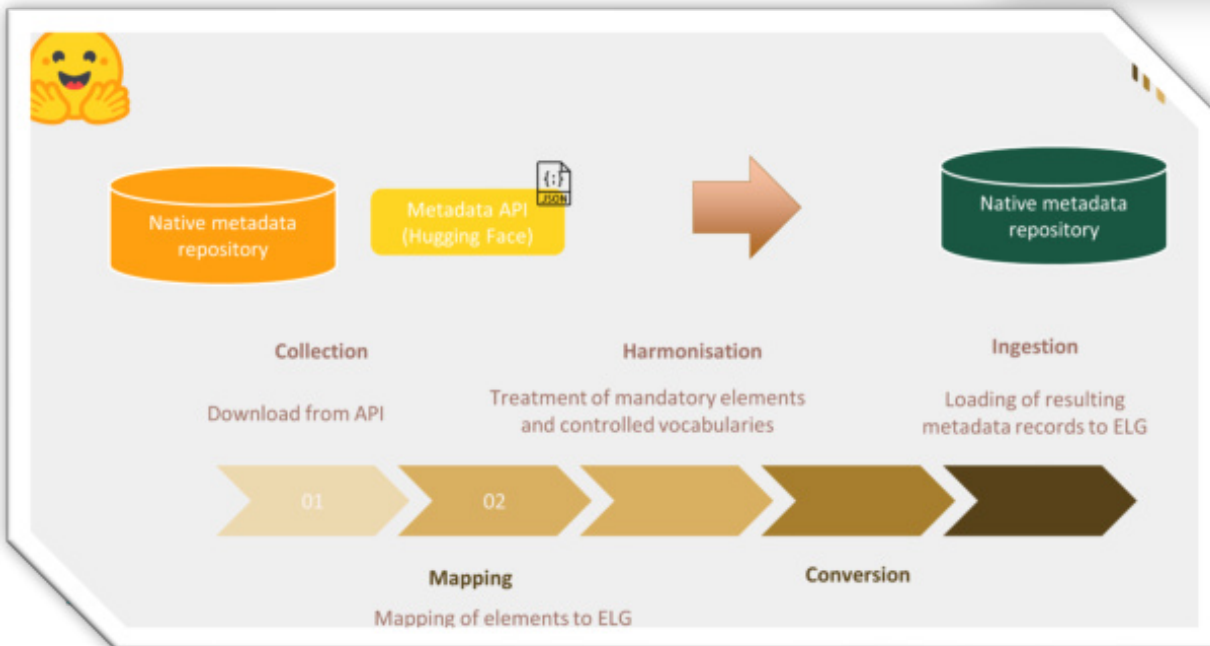
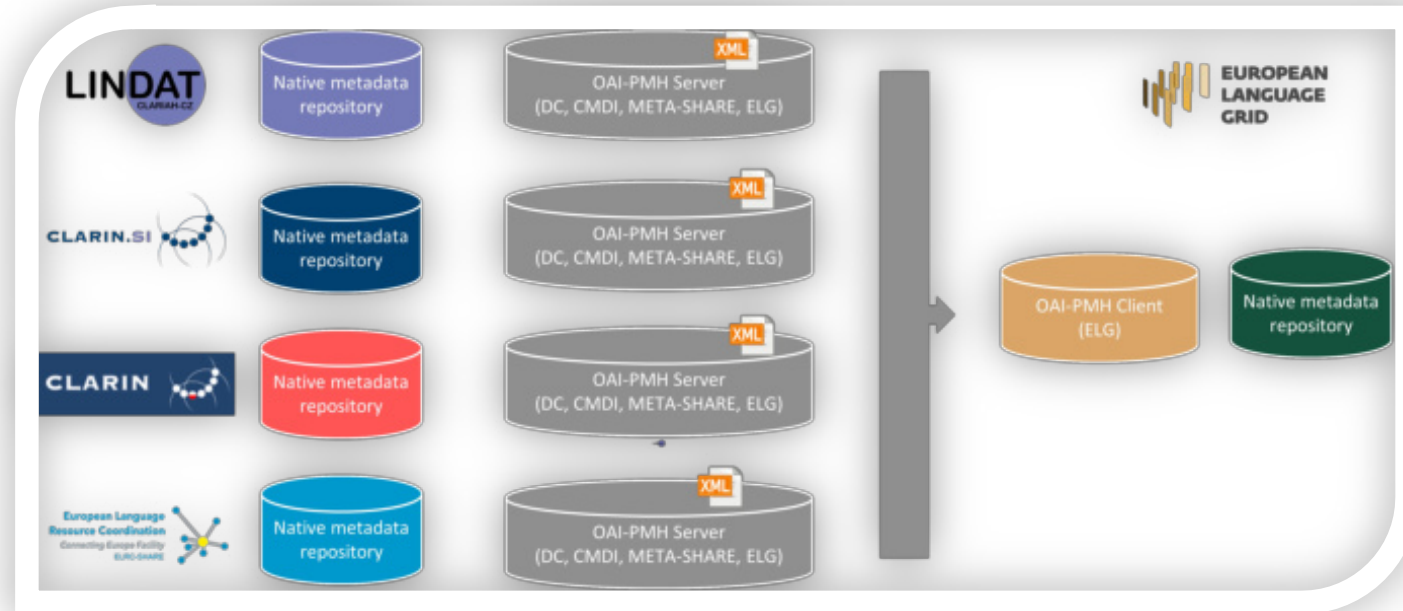




The ELG Catalogue

Continued population of ELG

Aggregating (meta)data
through multiple strategies



One Platform for all European Language Technologies

Discover, try out, use and download LT services and resources for all European languages.

Browse ELG and find the LT services, resources, developers and providers you are looking for.

Search the catalogue

Search

ELG RELEASE 30-JUNE 2022 — GRID WORKING — NO MAINTENANCE SCHEDULED

6257

Corpora

3175

Tools & Services

2249

Conceptual Resources

464

Models & Grammars

1779

Organizations

37

Projects

Browse ELG to find the resource you are looking for

The European Language Grid provides access to Language Technology resources from all over Europe. ELG contains tools and services, language resources and information on European LT companies and research organisations as well as their projects.

You can find out the complete current and past history of the LT

Do Europe's languages receive the same level of technology support? Check out our dashboard on Digital Language Equality to find out.

Visit the dashboard

One Platform for all European Language Technologies



EUROPEAN
LANGUAGE
GRID
RELEASE 1

Discover, try out, use and download LT
services and resources for all European
languages.

Browse ELG and find the LT services,
resources, developers and providers you
are looking for.

Search the catalogue

Search

Catalogue Documentation & Media About

June 2022

6257
Corpora

3175
Tools & Services

2249
Conceptual Resources

464
Models & Grammars

1779
Organizations

37
Projects

Discover, try out, use and download LT
services and resources for all European
languages.

Browse ELG and find the LT services,
resources, developers and providers you
are looking for.

Search the catalogue

Search

ELG RELEASE 3.0 (JUNE 2023) — GRID WORKING — NO MAINTENANCE SCHEDULED

8003
Corpora

3871
Tools & Services

2818
Conceptual Resources

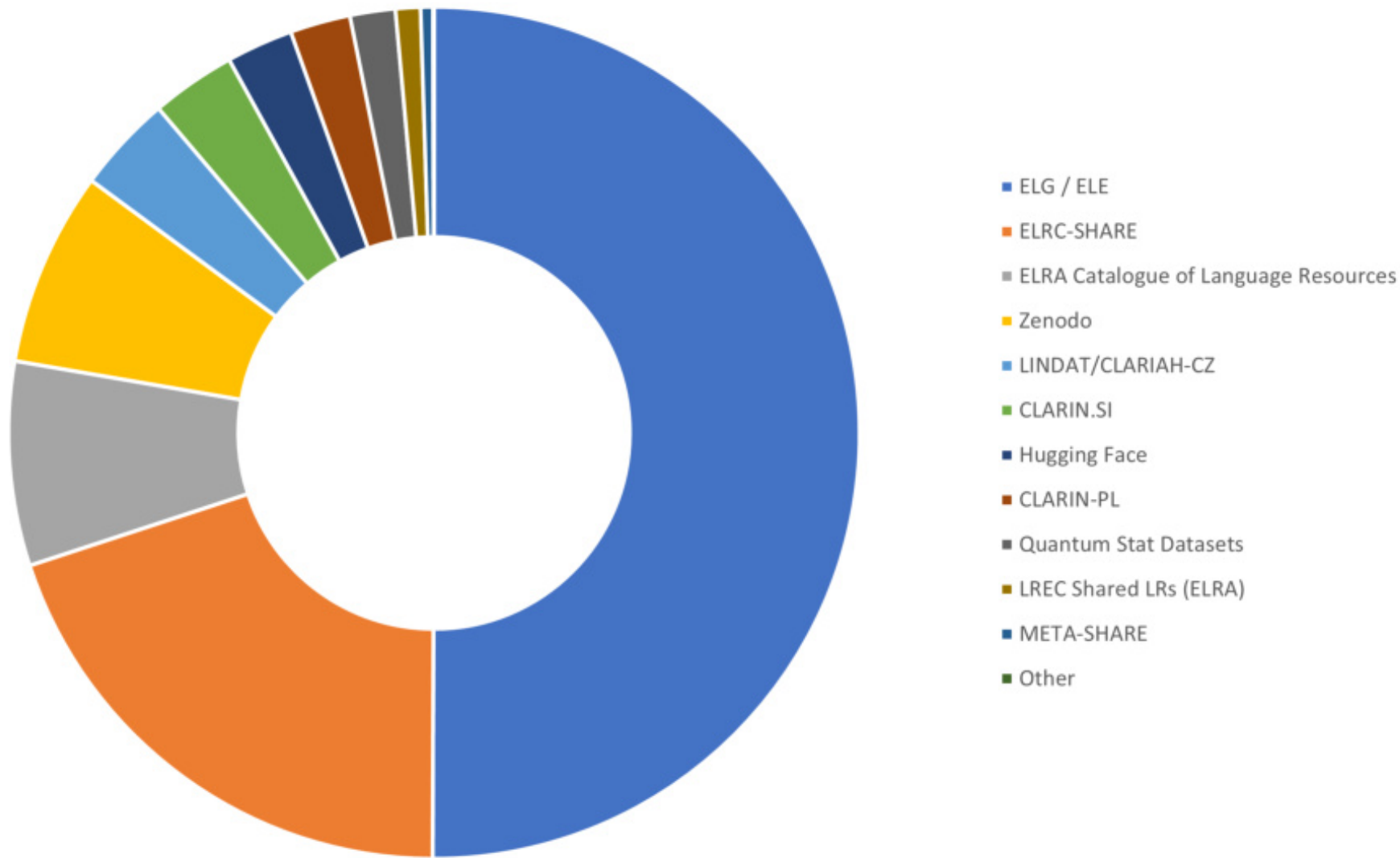
524
Models & Grammars

1774
Organizations

513
Projects

June 2023

Sources of Metadata Records





The DLE Dashboard

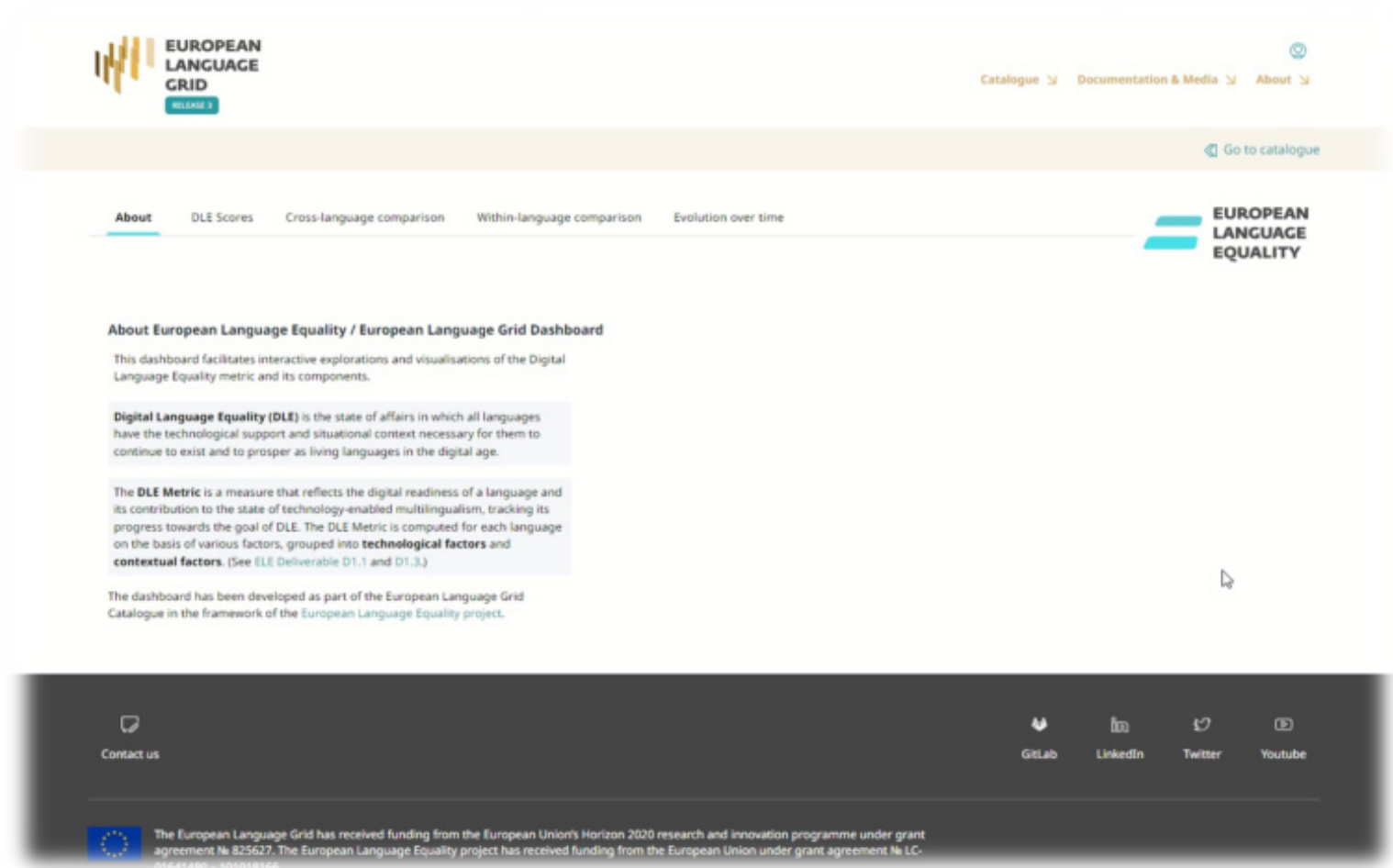
The DLE Dashboard

- A web-based interactive dashboard developed by ELE as part of the ELG Platform which exposes the contents of the ELG database as interactive visuals dynamically created by user queries
- It provides
 - an easy-to-use mechanism for exposing and monitoring the technological and contextual factors that contribute to DLE
 - an updatable and comparable measurement of the level of LT support and provision that each European language enjoys

<https://live.european-language-grid.eu/catalogue/dashboard>

Dashboard Walkthrough

- DLE technological and contextual scores
- Cross-language comparisons
 - As per resource types and features
 - Through
 - Histograms
 - Heatmaps
 - Radial bars
- Within-language comparisons
- Evolution of resources over time





Cross-Language Comparison

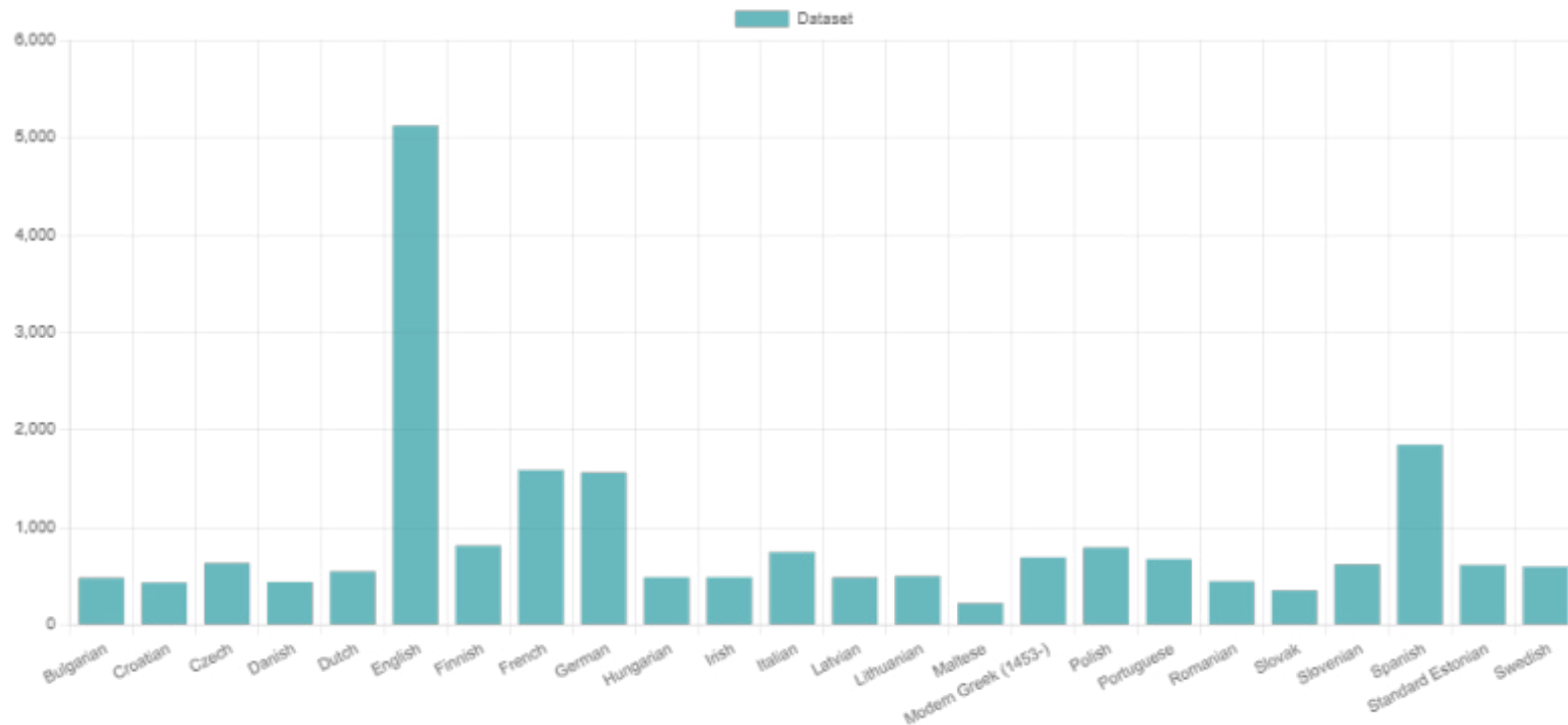
Language Resources and Technologies

- Corpora and lexical resources are the most numerous resource types, across all languages
- All languages seem to be better supported with translation technologies and text processing tools among all tools and services functions
- The least populated technologies are HCI, NLG and image/video processing

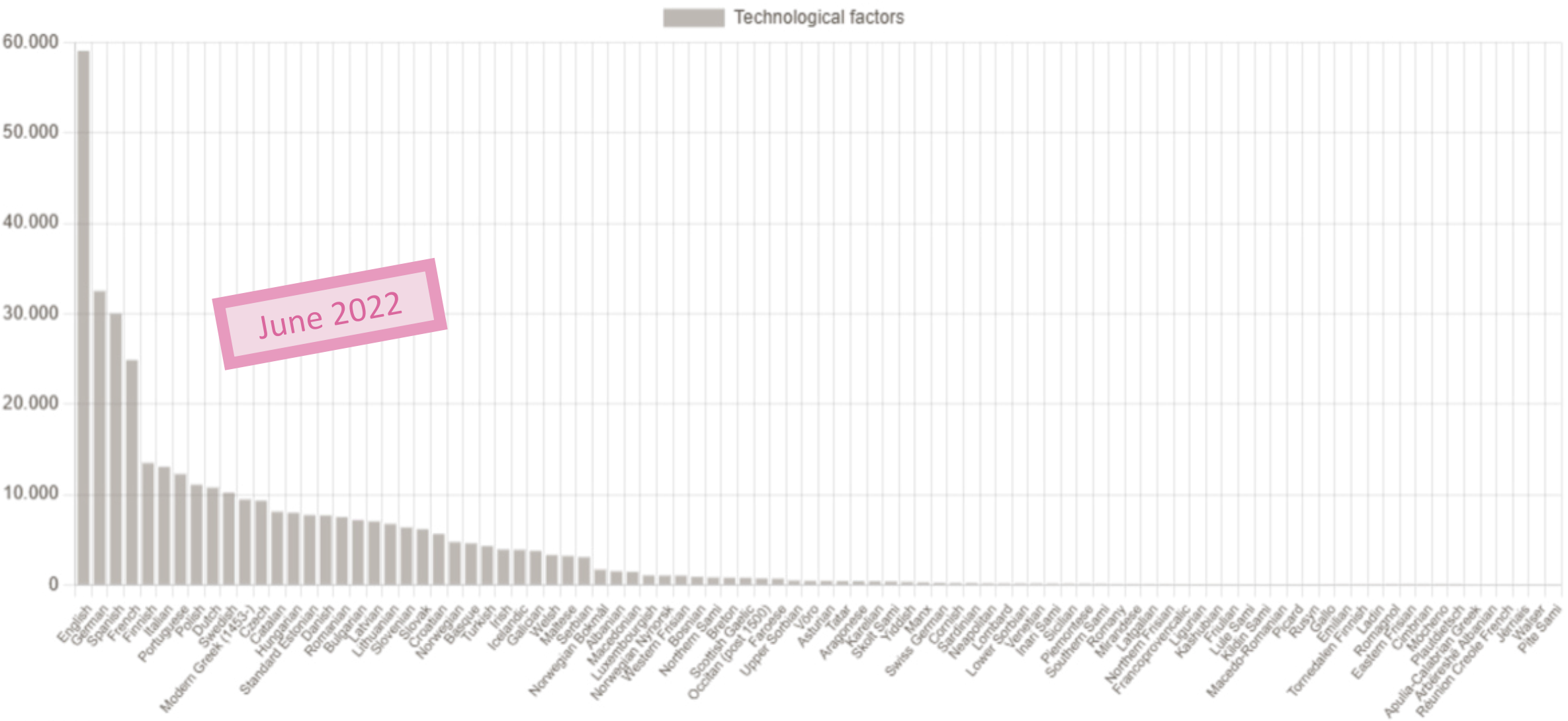


Language Resources and Technologies

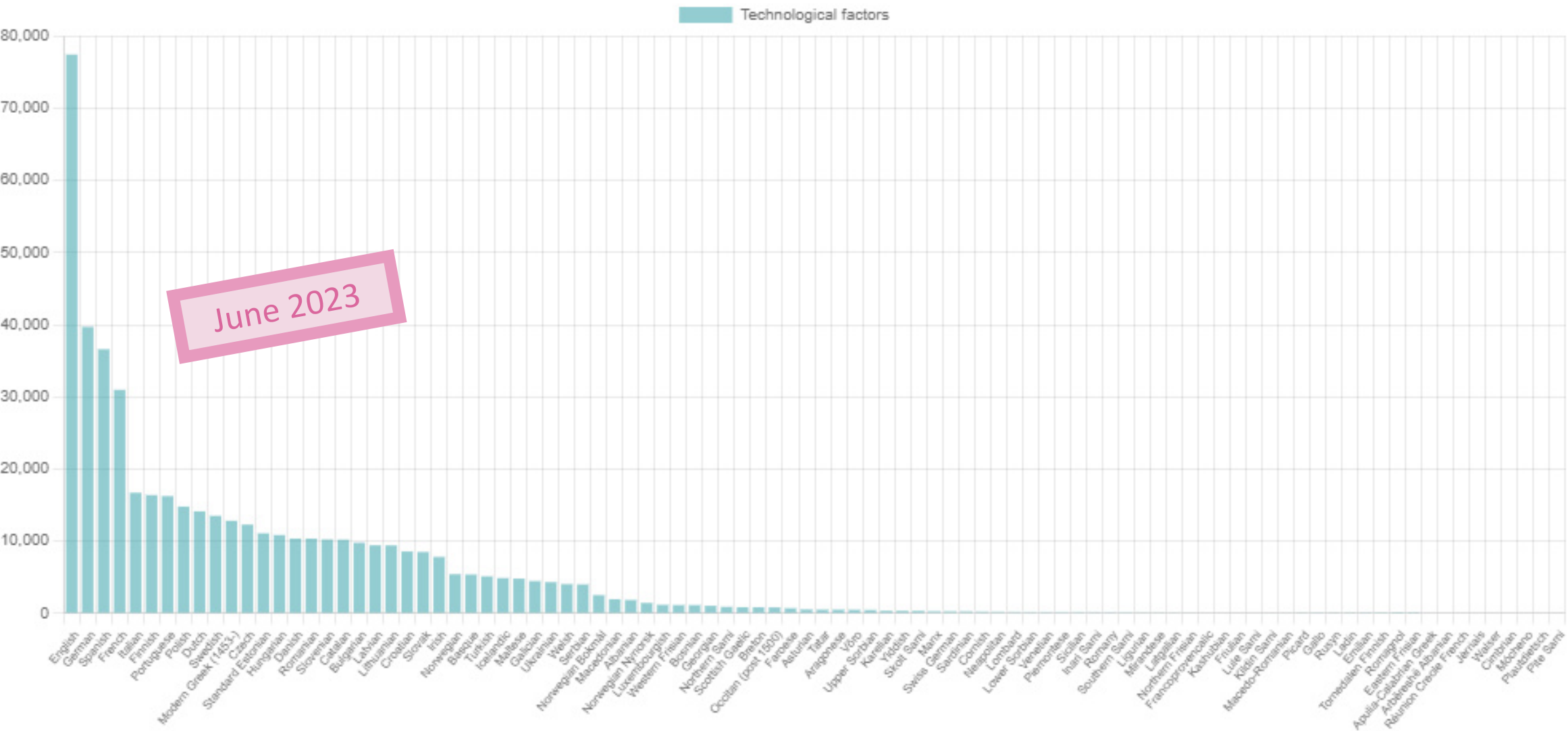
- Multimodal datasets (media type: audio, image or video): all languages are weakly covered



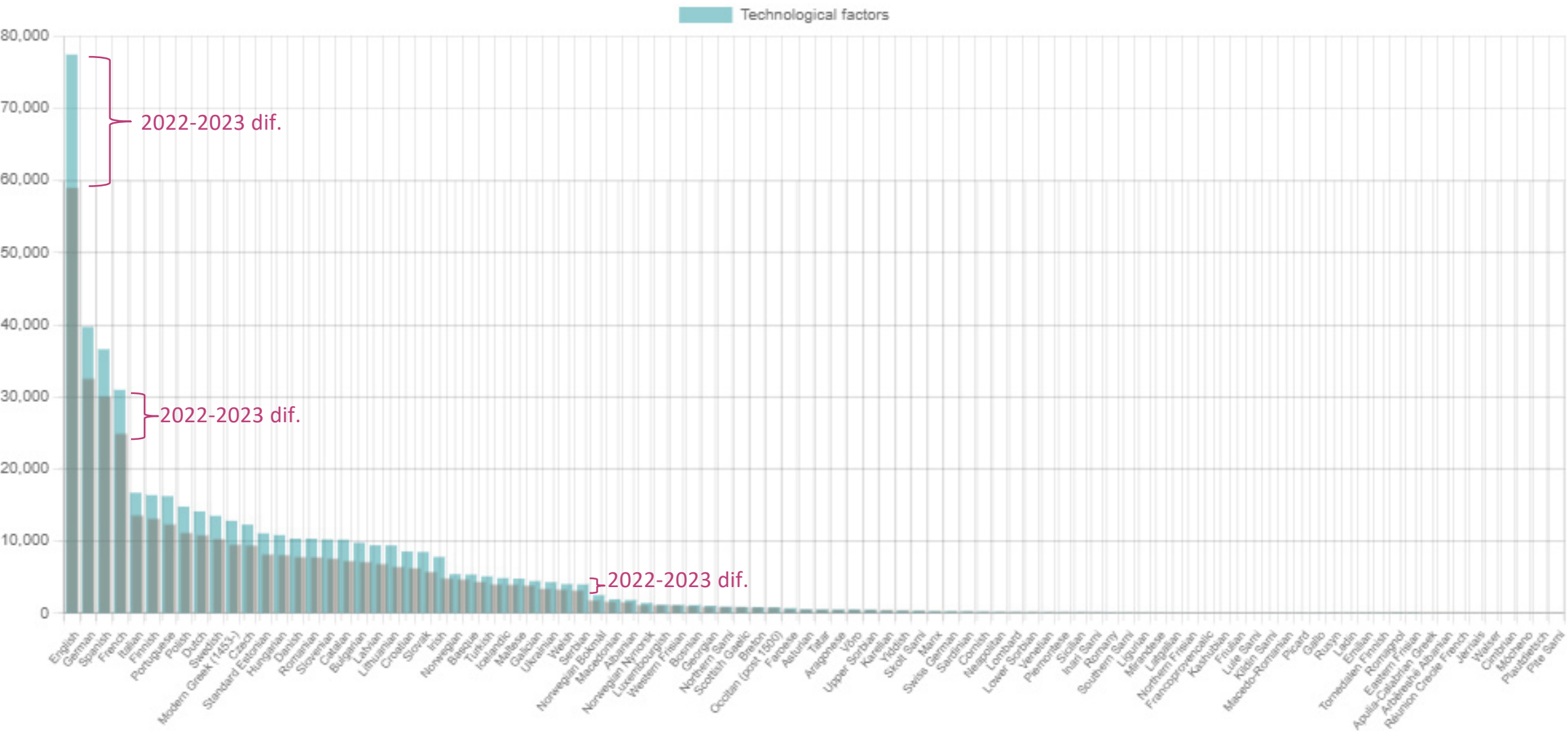
Technological DLE Scores (1/4)



Technological DLE Scores (2/4)



Technological DLE Scores (3/4)



Technological DLE Scores (4/4)

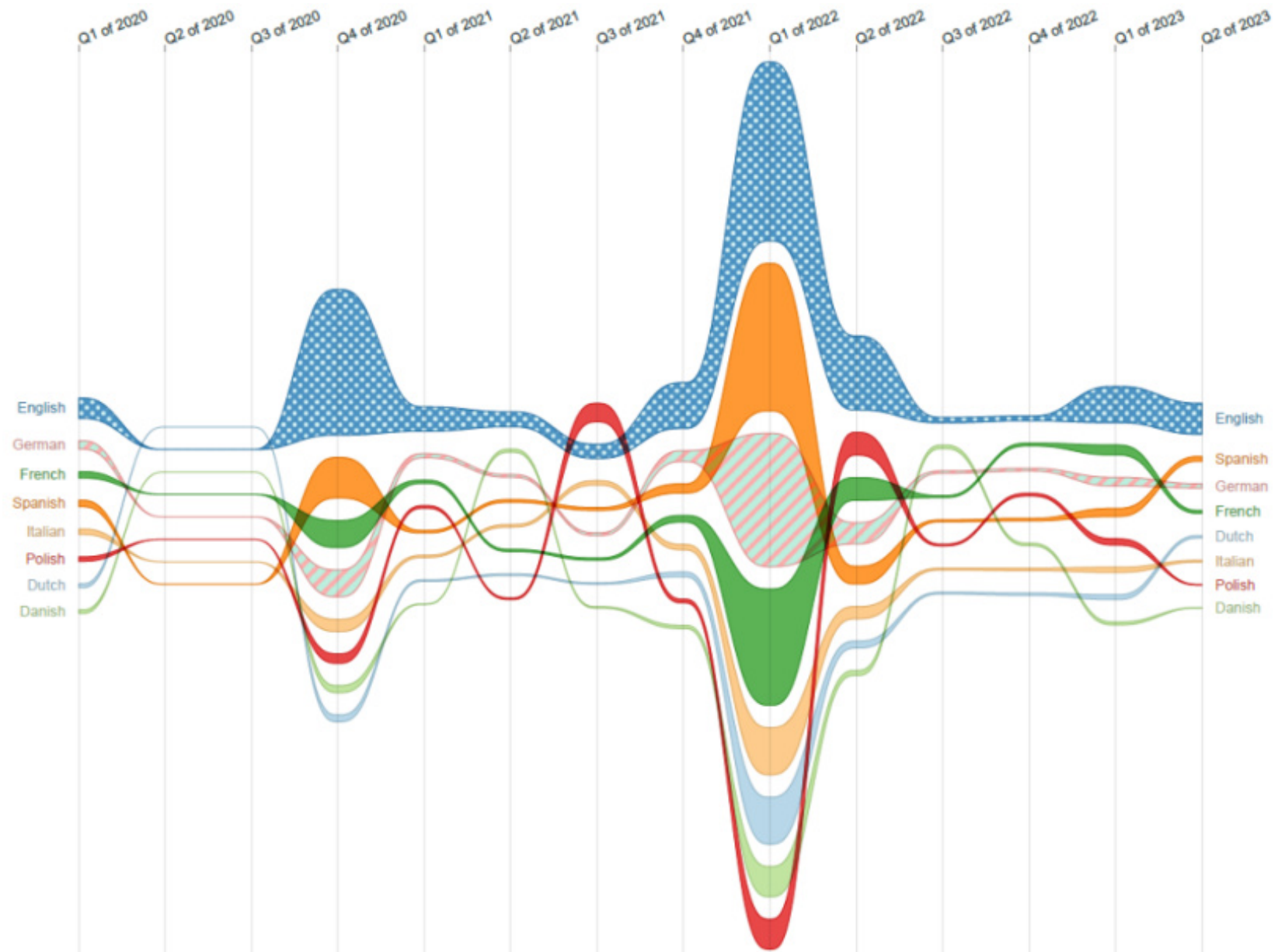
- Top-10 languages (all)

| 2022 | 2023 |
|------------|------------|
| English | English |
| German | German |
| Spanish | Spanish |
| French | French |
| Finnish | Italian |
| Italian | Finnish |
| Portuguese | Portuguese |
| Polish | Polish |
| Dutch | Dutch |
| Swedish | Swedish |

- Top-5 **non-EU official** languages

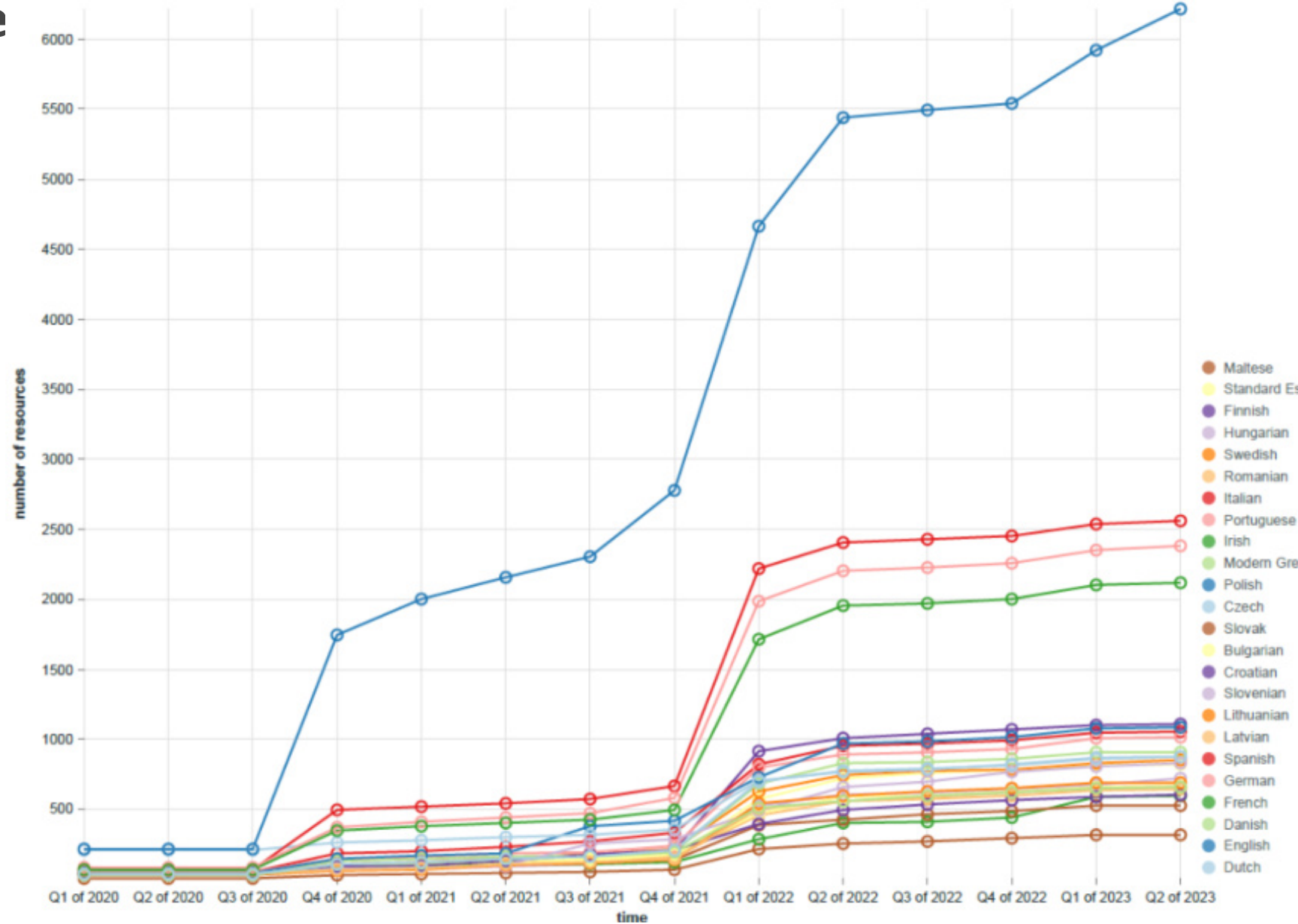
| 2022 | 2023 |
|-----------|-----------|
| Catalan | Catalan |
| Norwegian | Norwegian |
| Basque | Basque |
| Turkish | Turkish |
| Icelandic | Icelandic |

Evolution Over Time (1/2)



Evolution Over Time (2/2)

- All languages are progressing
- English shows steeper progress in Q1-Q2 2023
- Gap has increased





Thank you!



The European Language Equality project has received funding from the European Union under grant agreements № LC-01641480 – 101018166 (ELE) and № LC-01884166 – 101075356 (ELE2).

Maria Giagkou, Stelios Piperidis, Athanasia Kolovou, Miltos Deligiannis, Leon Voukoutis, Penny Labropoulou
(Institute for Language and Speech Processing / RC Athena)

27-06-2023 META-FORUM 2023 – Digital Language Equality for a Multilingual Europe
<http://european-language-equality.eu>