

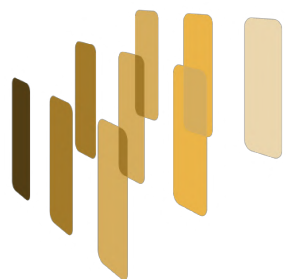
EUROPEAN LANGUAGE GRID



European Language Grid: General Overview – Platform – Services and Resources

Georg Rehm (DFKI), Stelios Piperidis (ILSP, R.C. “Athena”), Kalina Bontcheva (University of Sheffield),
georg.rehm@dfki.de, spip@athenarc.gr, k.bontcheva@sheffield.ac.uk

08/09-06-2022 META-FORUM 2022 – Joining the European Language Grid (hybrid conference)
<http://www.european-language-grid.eu> – <http://european-language-equality.eu>



EUROPEAN LANGUAGE GRID

Objectives (Selection)

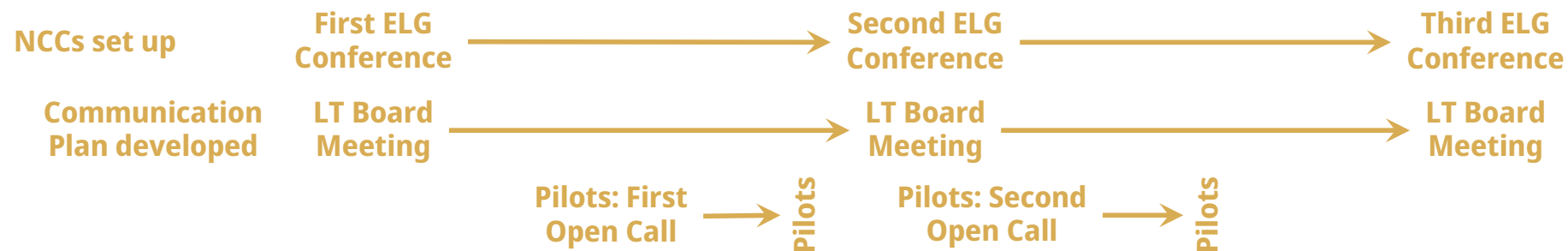
1. Establish the ELG as the primary Language Technology platform and market place in Europe to tackle the fragmentation of the European LT landscape.
2. ELG as a platform for commercial and non-commercial, industry-related LTs (functional and non-functional).
3. Enable the European LT community to upload services and data sets, to deploy them and to connect with, and make use of those resources made available by others.
4. Enable businesses to grow and benefit from scaling up.
5. Unleash enormous potential for innovation.



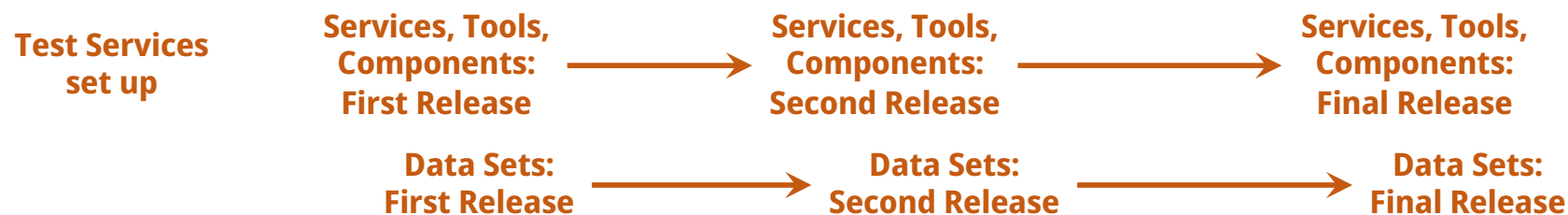
Kick-off meeting, 22/23 January 2019



Grid Community



Grid Content



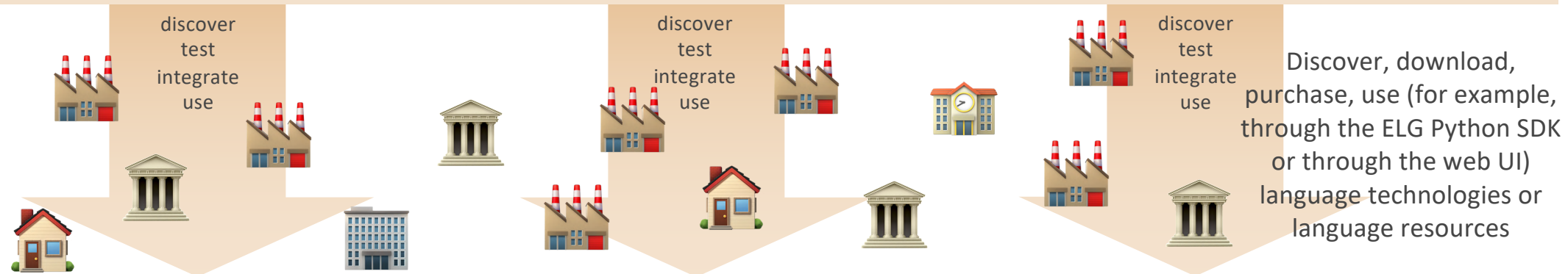
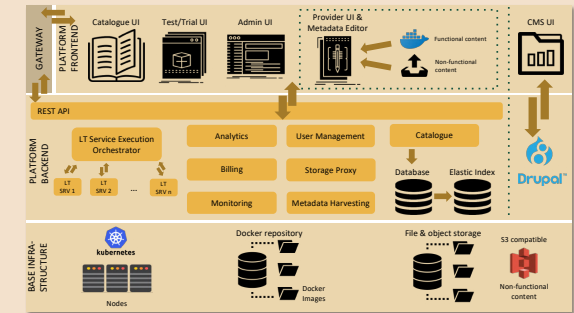
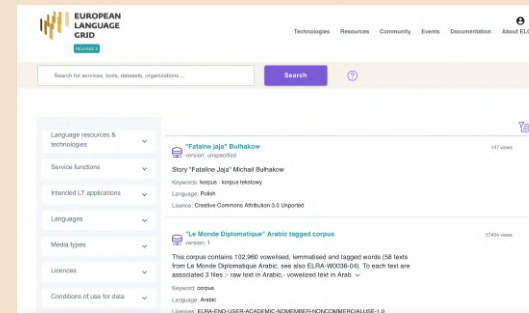
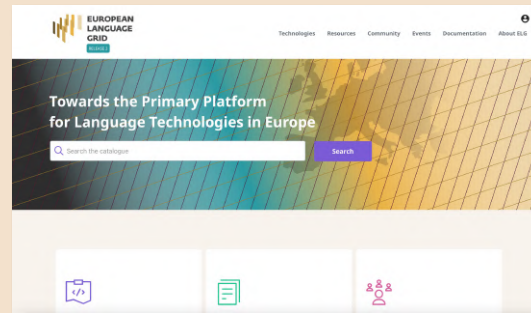
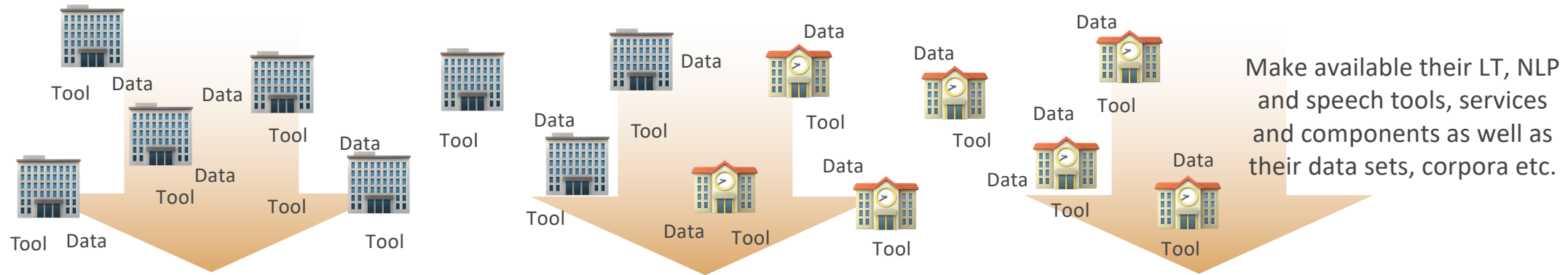
Grid Platform



Mgmt.

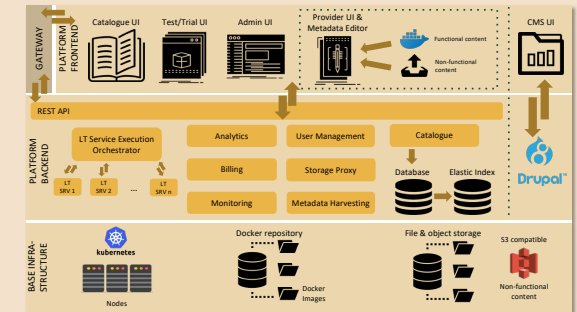
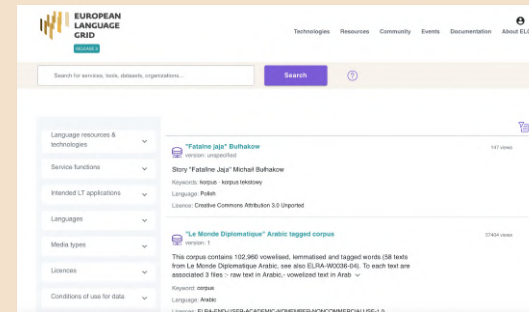
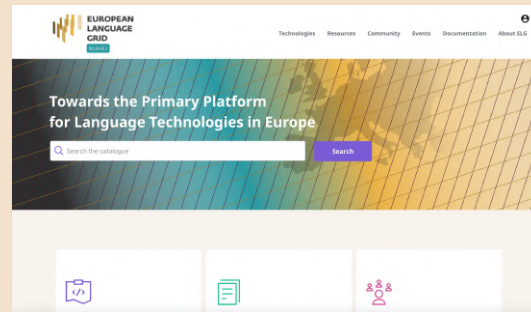


Developers of Language Technologies: Companies, Universities, Research Centres (approx. 1750-2000 organisations in total in Europe)

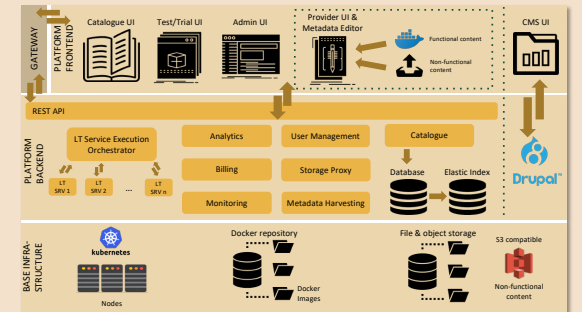
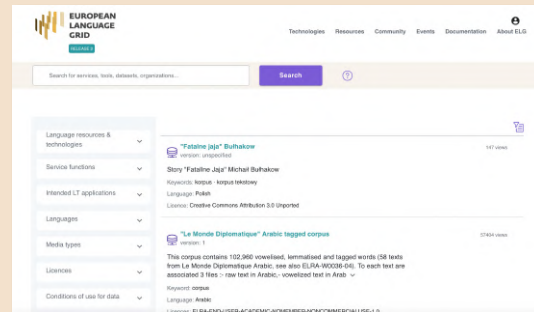
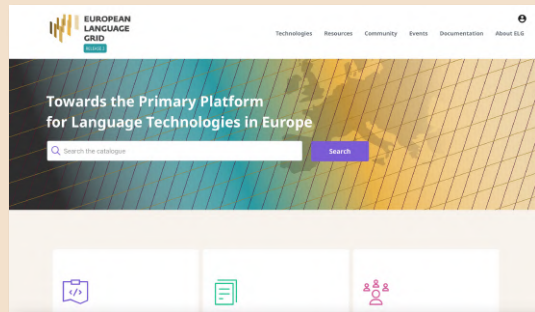




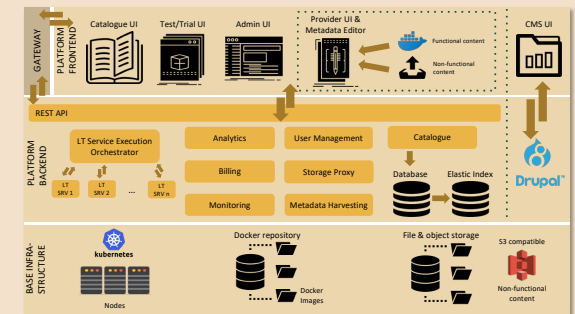
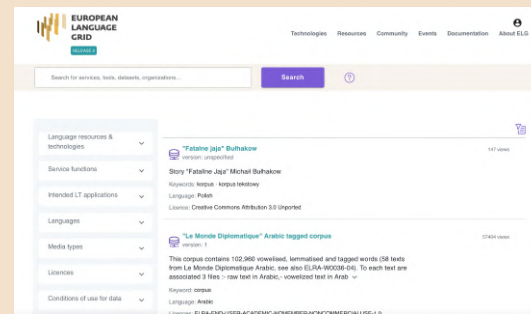
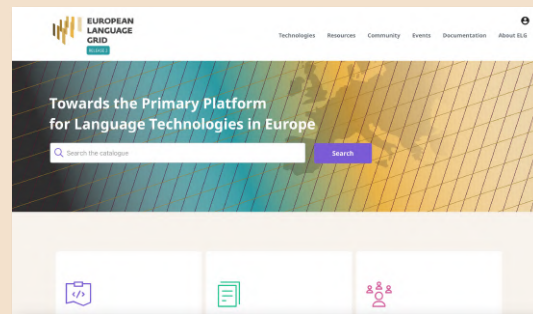
Harvesting of Language Technology-related Resource, Data, Tool and Model Repositories



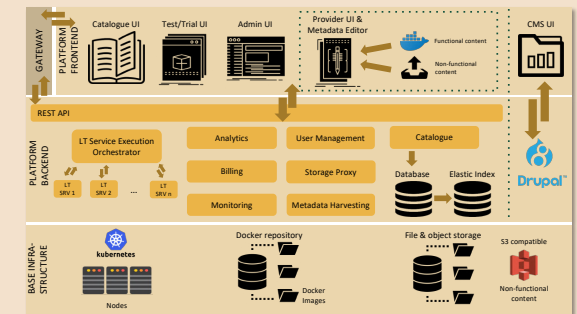
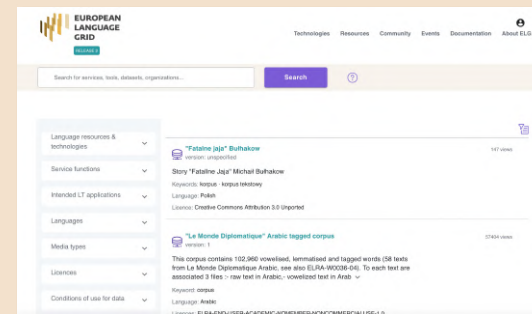
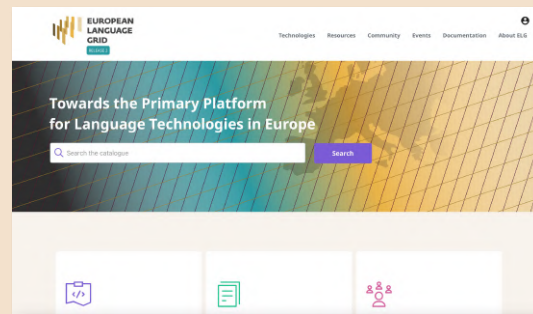
European Language Grid is a **joint technology platform** for the whole European Language Technology community (approx. 1750-2000 organisations)



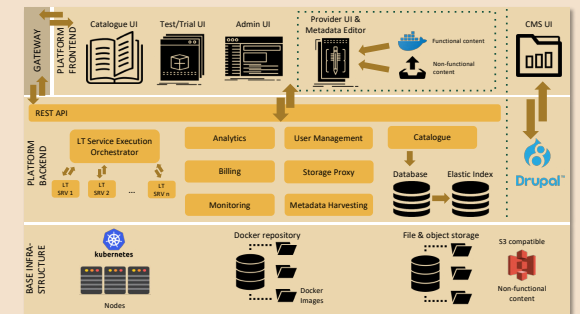
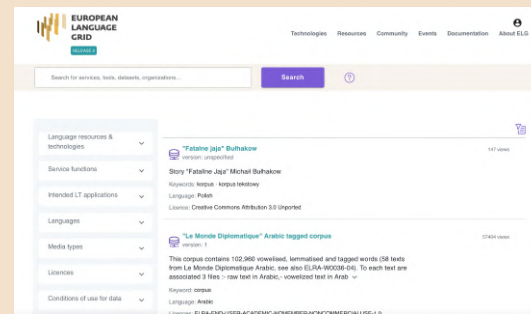
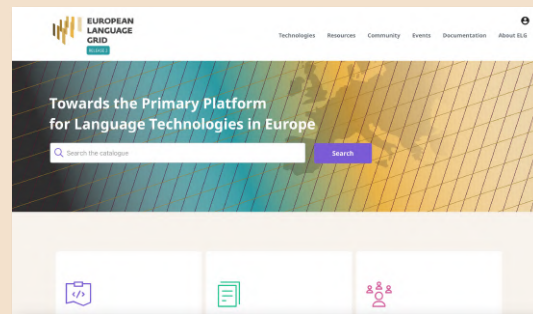
European Language Grid is a joint tool and resource sharing platform



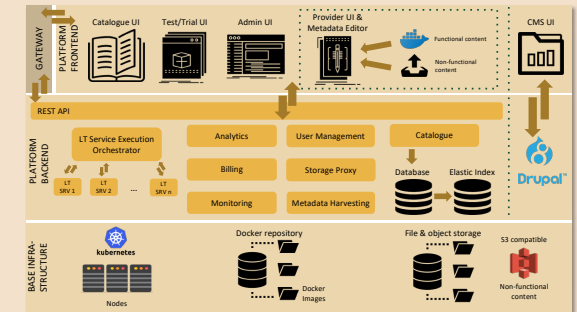
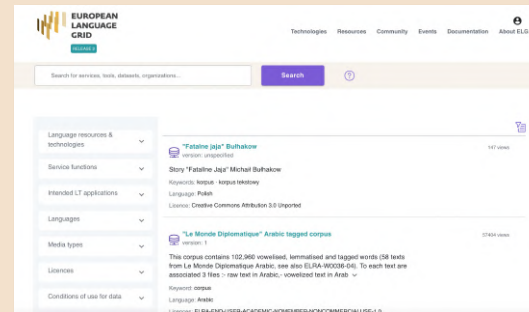
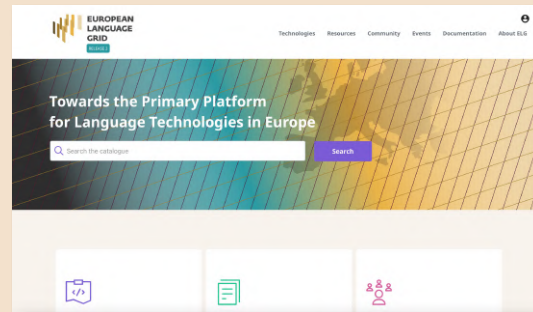
European Language Grid is a **language data space**



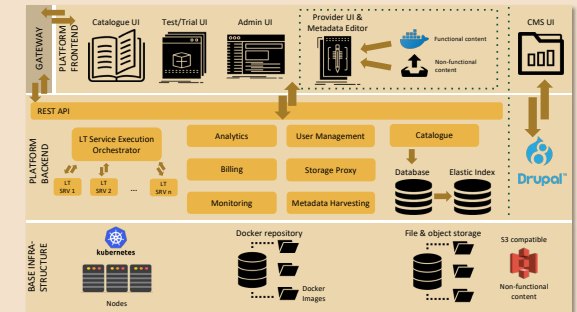
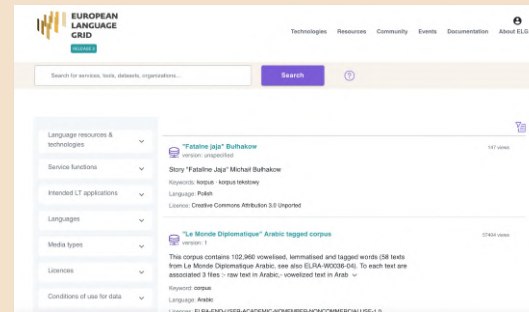
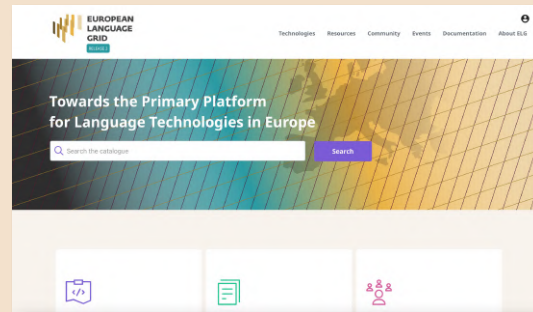
European Language Grid is a **marketplace** for the whole European Language Technology community



European Language Grid is the community of European Language Technology developers



European Language Grid is the **yellow pages** of the European Language Technology community



ELG makes available all European Language Technologies and Language Resources in a one-stop-shop.

ELG Release 3 (June 2022)

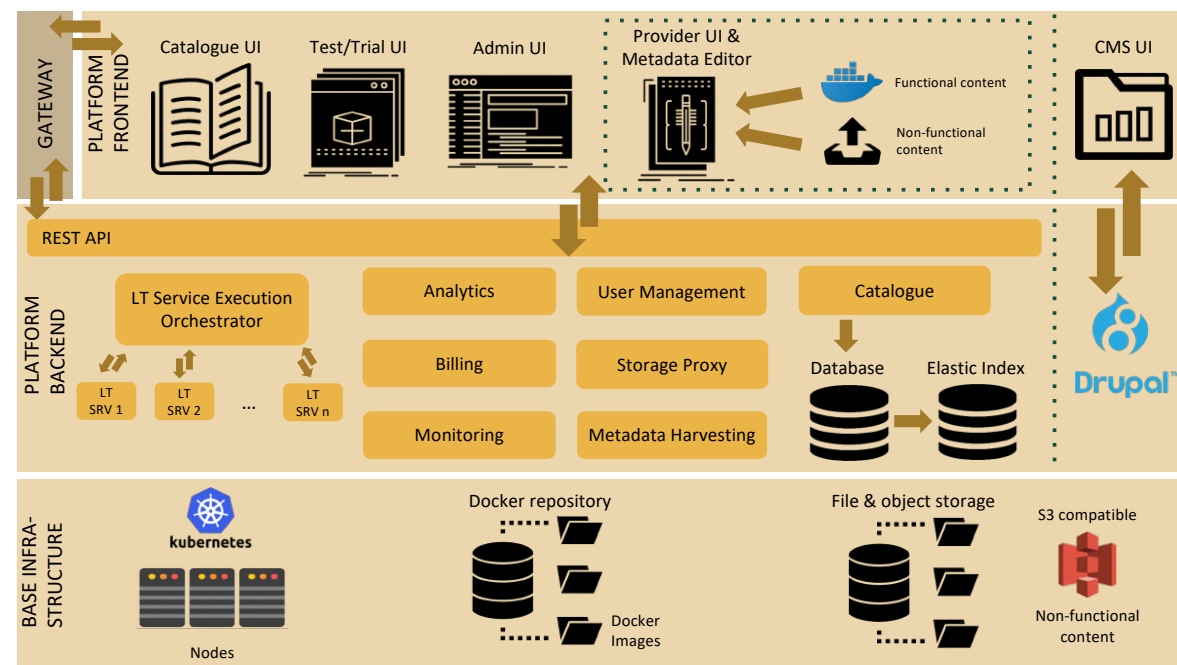
12,000+ Resources without organisations

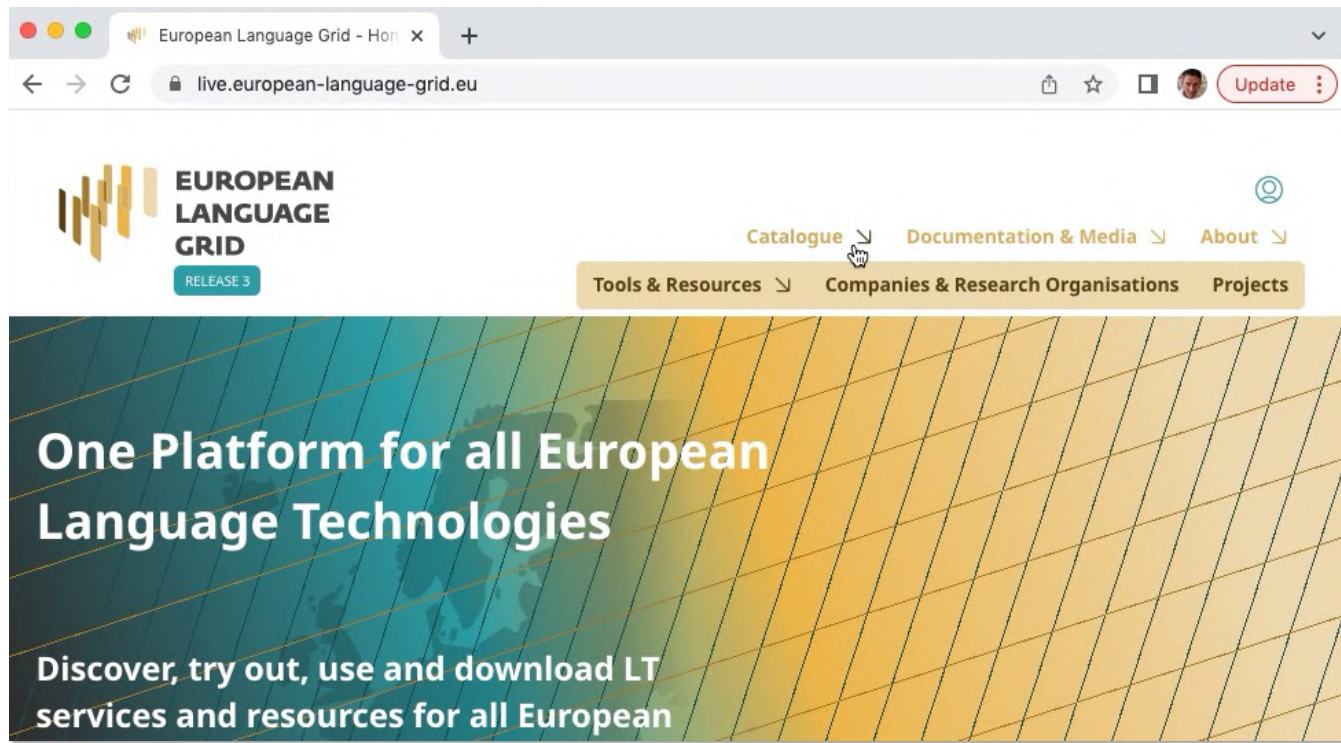
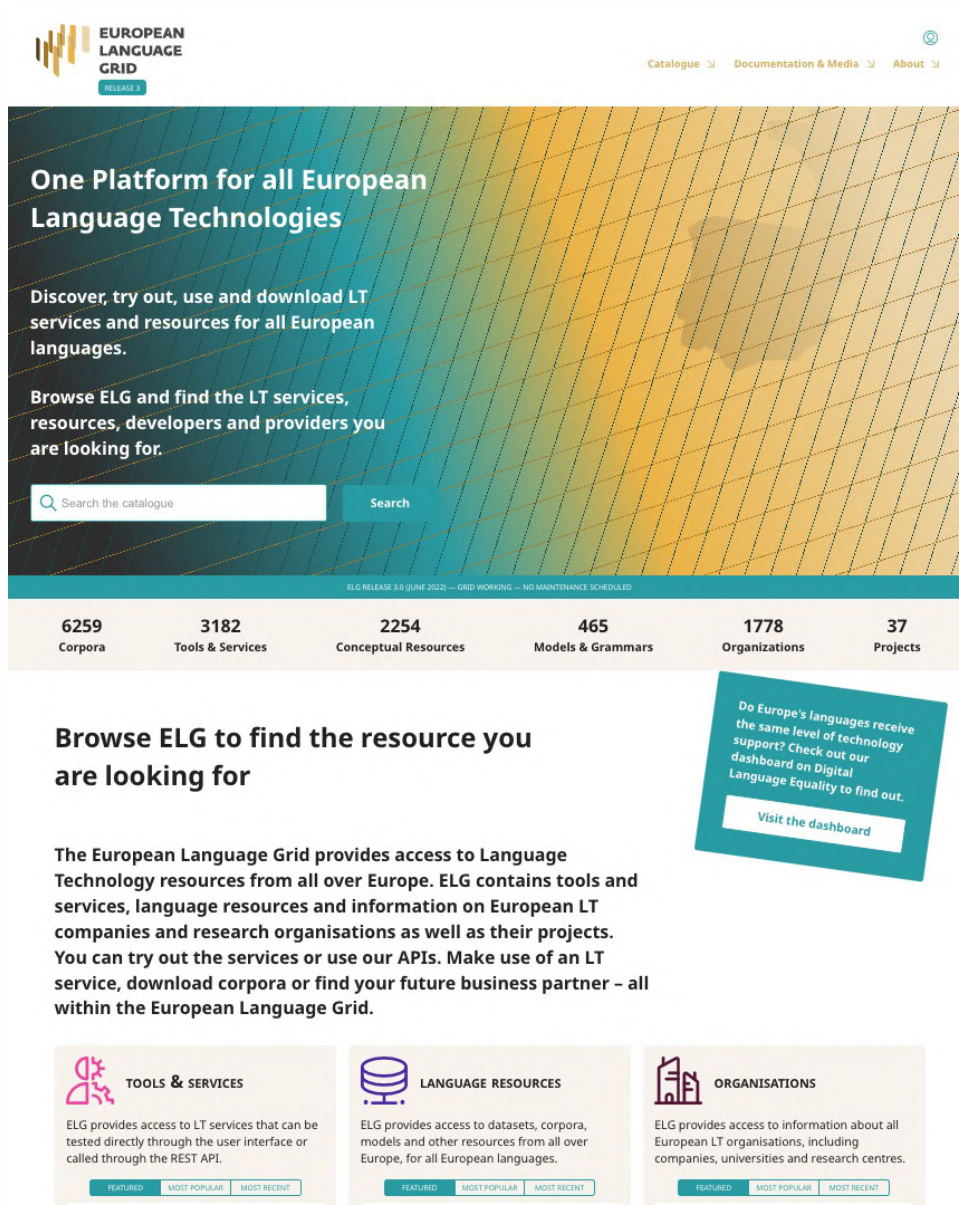
- 6259 corpora and data sets
- 3182 services and tools
 - 772 services available in and through ELG
- 2254 lexical/conceptual resources
- 465 models, grammars, lang. descriptions
- 1778 organisations (research orgs., companies)

Users can connect to the ELG cloud platform via ELG APIs, remote APIs, ELG GUI, Python SDK, download of containers or source code.

+ ELG / ELE	(9032)	+ CLARIN-PL	(360)
+ ELRC-SHARE	(1299)	+ Quantum Stat Datasets	(261)
+ ELRA Catalogue of Language Resources	(1180)	+ CLARIN.SI	(218)
+ Zenodo	(513)	+ LREC Shared LRs (ELRA)	(144)
+ LINDAT/CLARIAH-CZ	(507)	+ META-SHARE/ILSP	(69)
+ Hugging Face	(385)	+ META-SHARE/DFKI	(2)
		+ TUDatalib	(1)

Sources of data sets





List with 600+ projects
to be integrated soon

Stakeholders and Users

Companies that

- ... *develop* Language Technologies
- ... *integrate* Language Technologies
- ... *purchase* Language Technologies

Universities and research centres that

- ... *develop* Language Technologies
- ... *use* Language Technologies

Public administrations that *purchase* or *use* Language Technologies

Other organisations (e.g., NGOs) that *purchase* or *use* Language Technologies

Funding agencies that support the development of Language Technologies



META-FORUM 2019 (8/9 October) – Brussels, Belgium

Stakeholders and Users and Collaborators

32 ELG National Competence Centres

ELG Open Calls: 15 pilot projects

ELE – European Language Equality

AI4EU – European AI on demand platform

Data Spaces and Research Data Initiatives

ICT-29b research and innovation projects

Other related projects: CEF INEA projects (MAPA, NTEU, Microservices at your Service etc.), OpenGPT-X, NFDI4DS etc.

Additional initiatives: CLAIRE, AI PPP, ELRC, ECSPM, EFNIL, BDVA, CLARIN, W3C etc.



Open Calls for Pilot Projects

Two open calls for pilot projects

- Open Call #1: 03/04 2020
- Open Call #2: 10/11 2020

Pilot projects should or had to

- Type A: broaden ELG's portfolio or
- Type B: demonstrate usefulness of ELG

Up to €200,000 per project

Approx. €2,000,000 FSTP in total

Available in Open Call #1: 1.3M€

Available in Open Call #2: 585k€

Project duration: 9-12 months

Eligibility: SMEs, research organisations

Open Call #1 – Statistics

110 project proposals evaluated

- Type A: 79 – Type B: 31 proposals

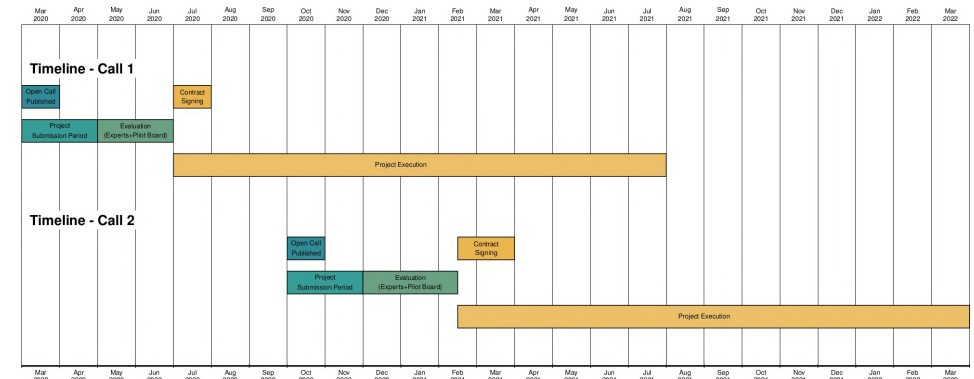
Total amount requested: 16.9M€

→ 10 projects selected

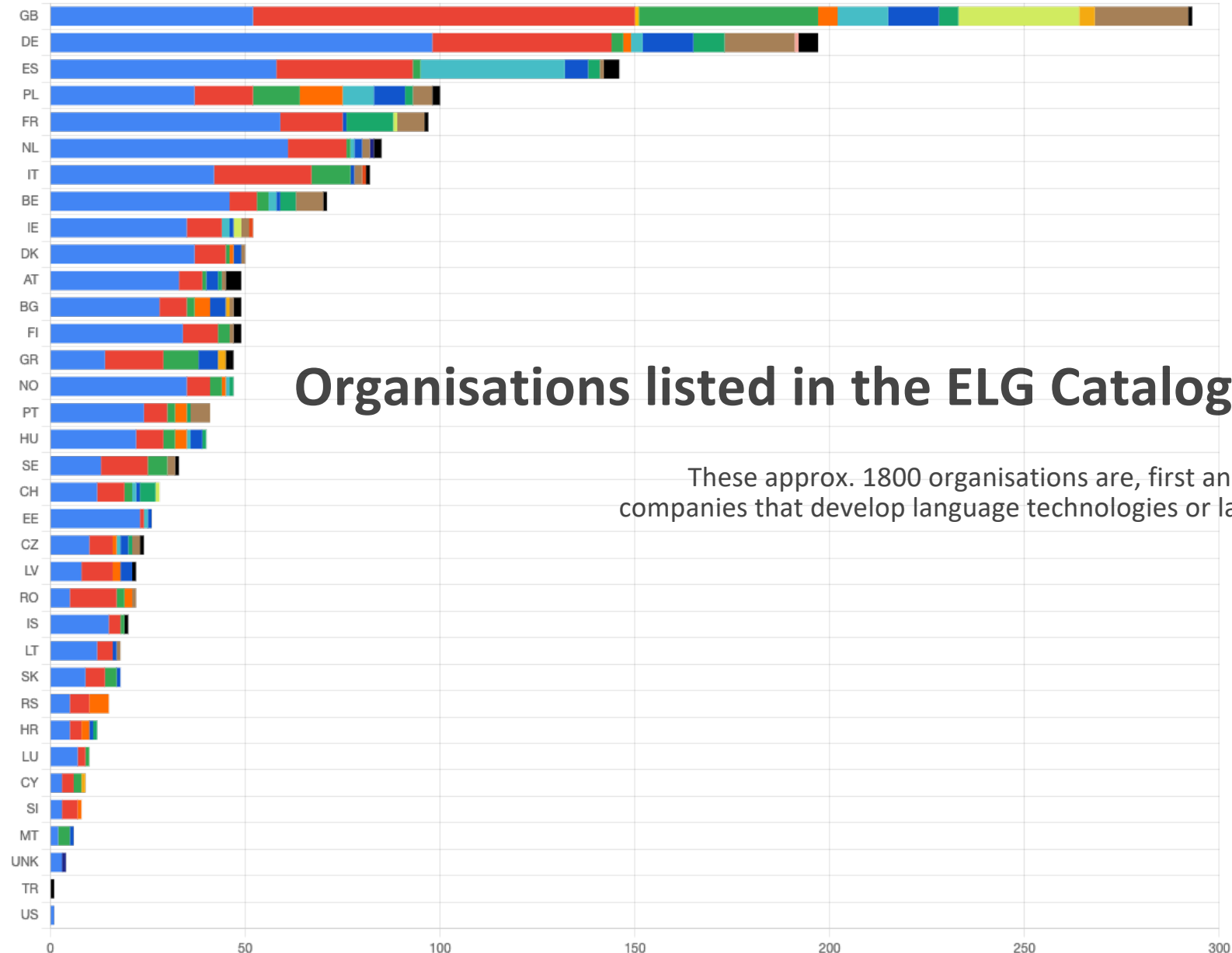
Open Call #2 – Statistics

106 project proposals evaluated

→ 5 projects selected



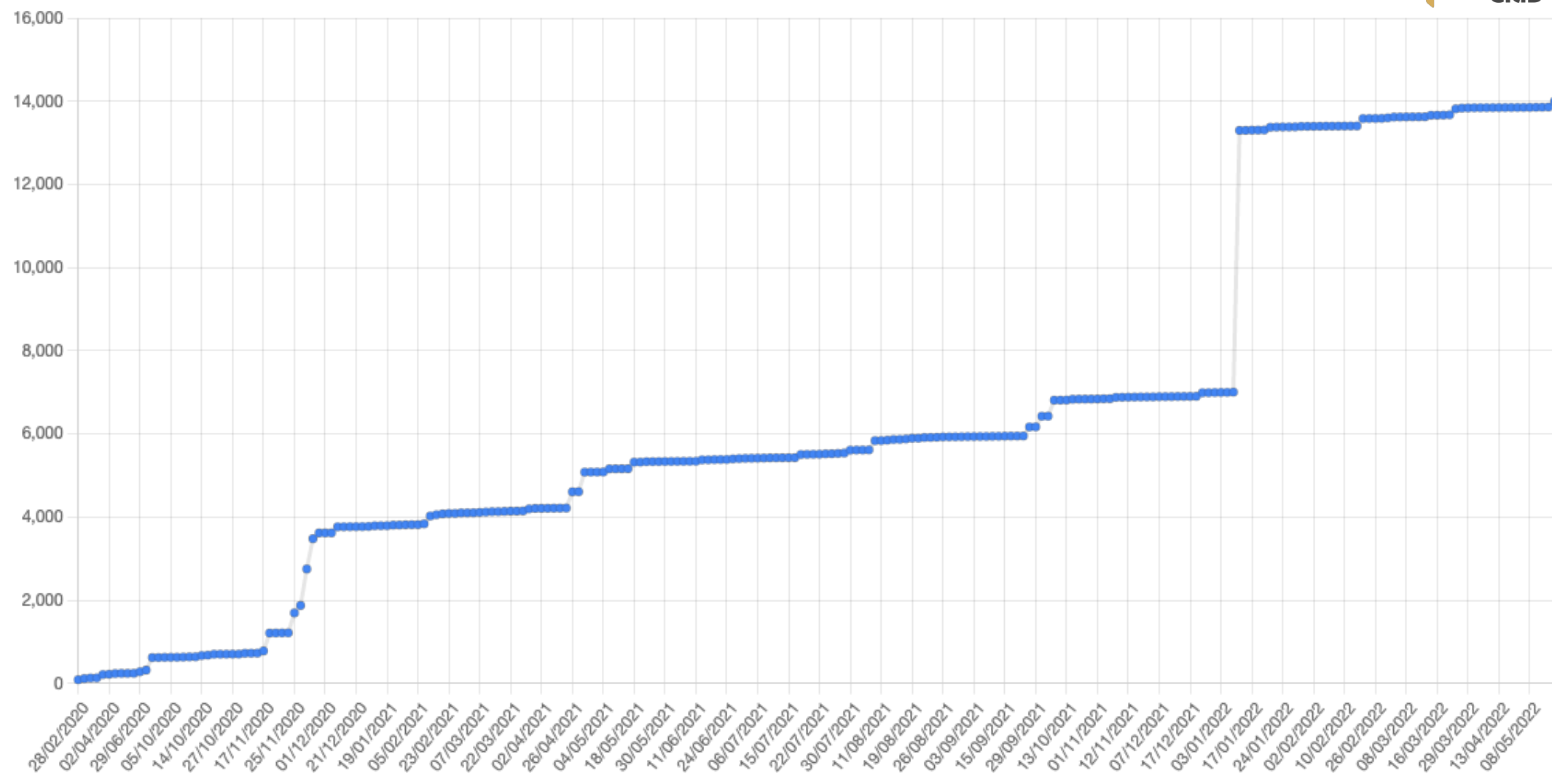
Open Call #1	Fondazione Bruno Kessler	European Clinical Case Corpus ✓	IT	139,370€
	Lingsoft, Inc.	Lingsoft Solutions as Distributable Containers ✓	FI	140,625€
	Coreon GmbH	MKS as Linguistic Linked Open Data ✓	DE	167,375€
	Elhuyar Fundazioa	Basque-speaking smart speaker based on Mycroft AI ✓	ES	117,117€
	Universita' Degli Studi di Torino	Italian EVALITA Benchmark Linguistic Resources, NLP Services and Tools ✓	IT	126,125€
	University of Helsinki	Open Translation Models, Tools and Services ✓	FI	154,636€
	University of Vienna	Extracting Terminological Concept Systems from Natural Language Text ✓	AT	132,977€
	University of Turku	Textual paraphrase dataset for deep language modelling ✓	FI	166,085€
	Weber Consulting KG	Virtual Personal Assistant Prototype ✓	AT	87,445€
Open Call #2	FZI Research Center for IT	Streaming Language Processing in Manufacturing ✓	DE	132,160€
	Institute for Bulgarian Language	Multilingual Image Corpus 2021 ✓	BG	110,960€
	EDIA BV	CEFR Labelling and Assessment Services ✓	NL	137,560€
	University of West Bohemia	Motion Capture 3D Sign Language Resources ✓	CZ	85,421€
	Sapienza University of Rome	Universal Semantic Annotator ✓	IT	113,228€
	Sign Time GmbH	Sign language explanations for terms in a text ✓	AT	137,227€



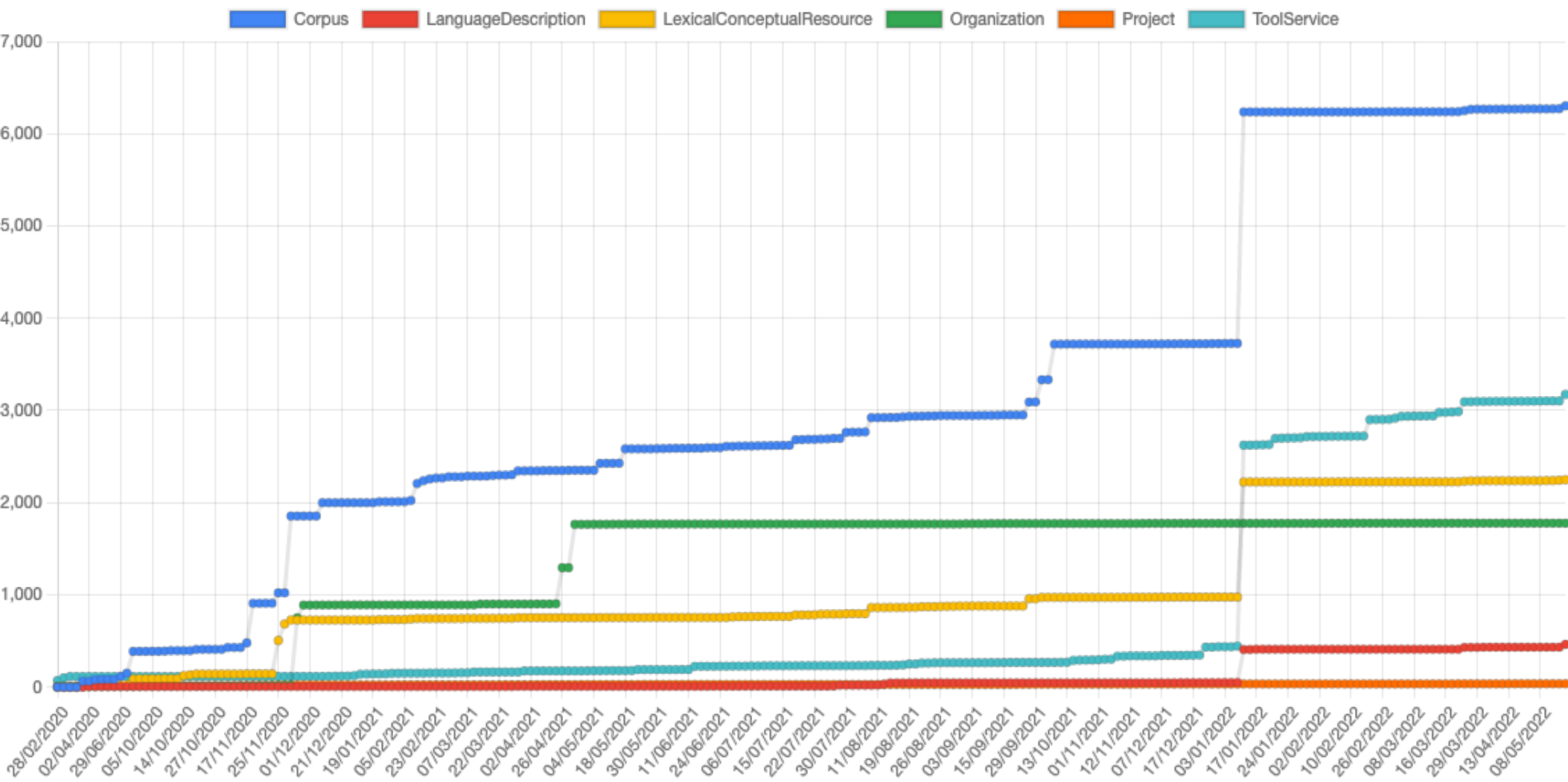
Organisations listed in the ELG Catalogue (by country)

These approx. 1800 organisations are, first and foremost, research groups or companies that develop language technologies or language resources (June 2022).

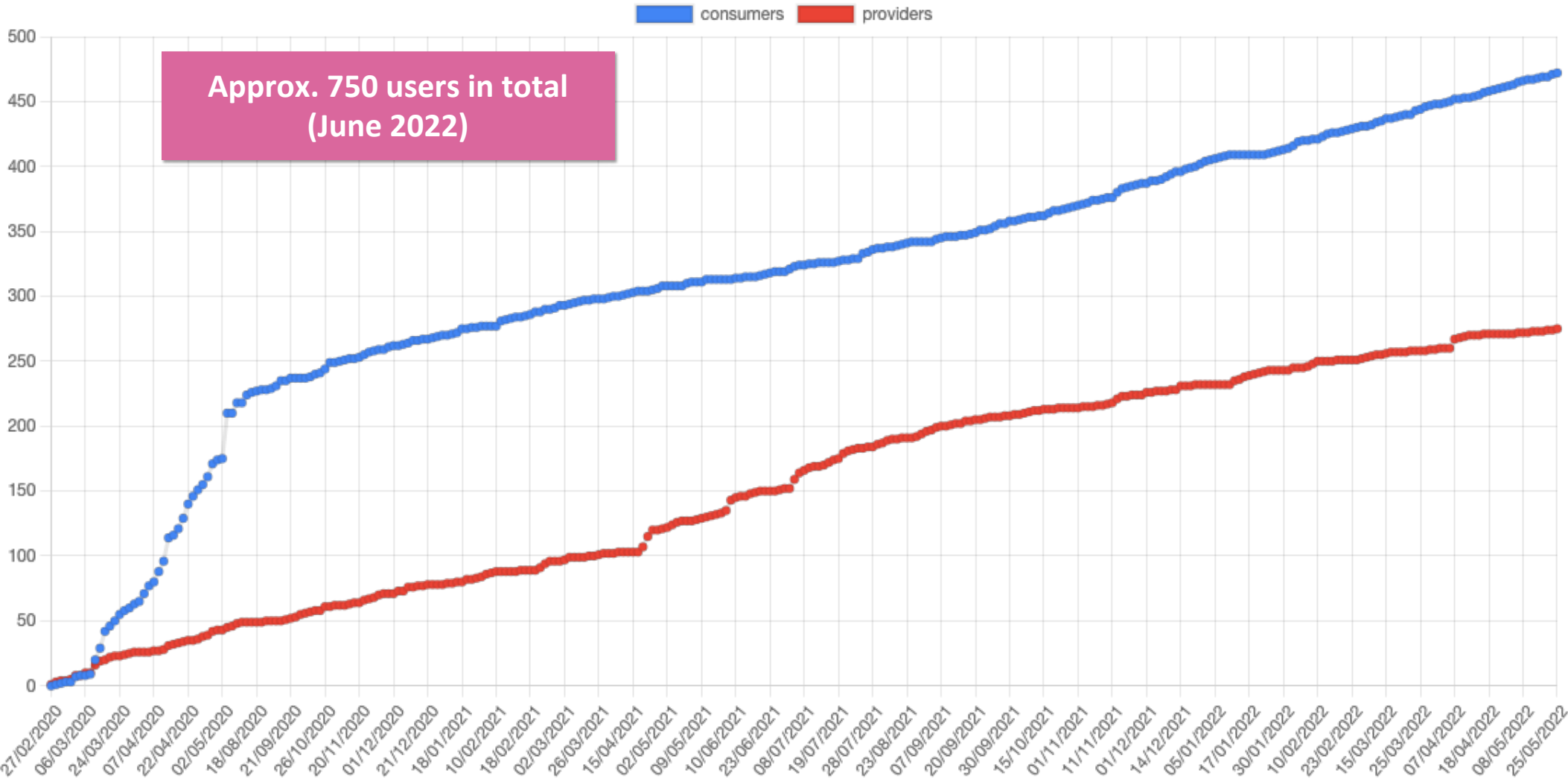
Number of Resources over Time



Number of Resources and Resource Types over Time

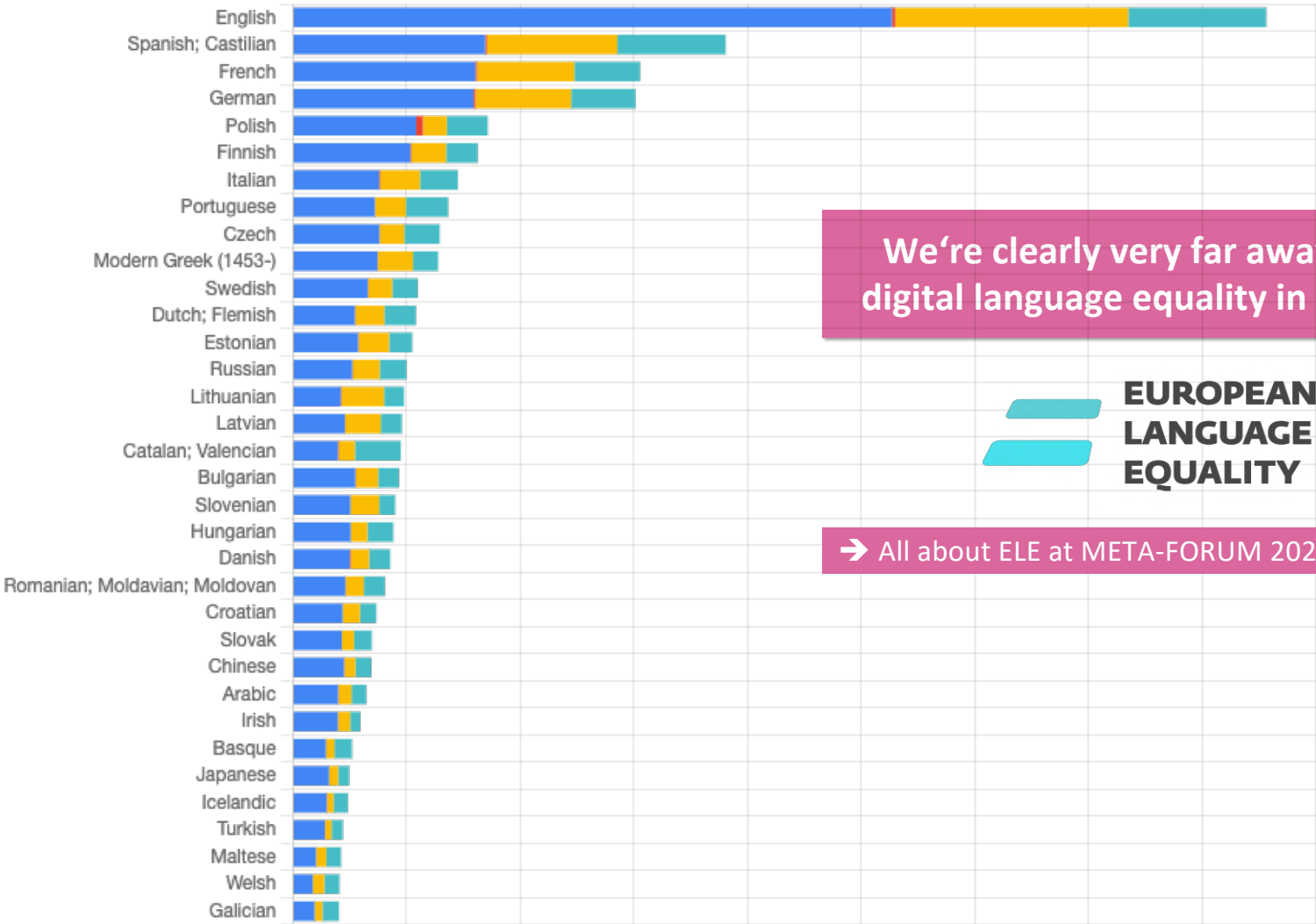


Number of Users (Consumers of LT, Providers of LT)



Languages of the Resources *(long tail cropped)*

Corpus LanguageDescription LexicalConceptualResource Organization Project ToolService



We're clearly very far away from digital language equality in Europe.

EUROPEAN
LANGUAGE
EQUALITY

→ All about ELE at META-FORUM 2022 tomorrow