

# EUROPEAN LANGUAGE EQUALITY

**D4.3**

**Report on  
EP/EC Workshop**

Authors	Georg Rehm, Stefanie Hegele, Katrin Marheinecke
Dissemination level	Public
Date	30-06-2022

## About this document

Project	European Language Equality (ELE)
Grant agreement no.	LC-01641480 – 101018166 ELE
Coordinator	Prof. Dr. Andy Way (DCU)
Co-coordinator	Prof. Dr. Georg Rehm (DFKI)
Start date, duration	01-01-2021, 18 months
Deliverable number	D4.3
Deliverable title	Report on EP/EC Workshop
Type	Report
Number of pages	21
Status and version	Final
Dissemination level	Public
Date of delivery	Contractual: 30-06-2022 – Actual: 30-06-2022
Work package	WP4: Communication – Dissemination – Exploitation – Sustainability
Task	Task 4.2 Liaise with EP/EC and organisation of a targeted workshop
Authors	Georg Rehm, Stefanie Hegele, Katrin Marheinecke
Reviewers	Aritz Farwell, German Rigau (UPV/EHU), Jane Dunne (DCU)
EC project officers	Susan Fraser, Miklos Druskoczi
Contact	<p>European Language Equality (ELE)  ADAPT Centre, Dublin City University  Glasnevin, Dublin 9, Ireland</p> <p>Prof. Dr. Andy Way – andy.way@adaptcentre.ie</p> <p>European Language Equality (ELE)  DFKI GmbH  Alt-Moabit 91c, 10559 Berlin, Germany</p> <p>Prof. Dr. Georg Rehm – georg.rehm@dfki.de  <a href="http://www.european-language-equality.eu">http://www.european-language-equality.eu</a>  © 2022 ELE Consortium</p>

## Consortium

1	Dublin City University (Coordinator)	DCU	IE
2	Deutsches Forschungszentrum für Künstliche Intelligenz GmbH (Co-coordinator)	DFKI	DE
3	Univerzita Karlova (Charles University)	CUNI	CZ
4	Athina-Erevnitiko Kentro Kainotomias Stis Pliroforias, Ton Epikoinonion Kai Tis Gnosis	ILSP	GR
5	Universidad Del Pais Vasco/ Euskal Herriko Unibertsitatea (University of the Basque Country)	UPV/EHU	ES
6	CROSSLANG NV	CRSLNG	BE
7	European Federation of National Institutes for Language	EFNIL	LU
8	Réseau européen pour l'égalité des langues (European Language Equality Network)	ELEN	FR
9	European Civil Society Platform for Multilingualism	ECSPM	DK
10	CLARIN ERIC – Common Language Resources and Technology Infrastructure as a European Research Infrastructure Consortium	CLARIN	NL
11	Universiteit Leiden (University of Leiden)	ULEI	NL
12	Eurescom (European Institute for Research and Strategic Studies in Telecommunications GmbH)	ERSCM	DE
13	Stichting LIBER (Association of European Research Libraries)	LIBER	NL
14	Wikimedia Deutschland (Gesellschaft zur Förderung freien Wissens e. V.)	WMD	DE
15	Tilde SIA	TILDE	LV
16	Evaluations and Language Resources Distribution Agency	ELDA	FR
17	Expert System Iberia SL	EXPSYS	ES
18	HENSOLDT Analytics GmbH	HENS	AT
19	Xcelerator Machine Translations Ltd. (KantanMT)	KNTN	IE
20	PANGEANIC-B. I. Europa SLU	PAN	ES
21	Semantic Web Company GmbH	SWC	AT
22	SIRMA AI EAD (Ontotext)	ONTO	BG
23	SAP SE	SAP	DE
24	Universität Wien (University of Vienna)	UVIE	AT
25	Universiteit Antwerpen (University of Antwerp)	UANTW	BE
26	Institute for Bulgarian Language “Prof. Lyubomir Andreychin”	IBL	BG
27	Sveučilište u Zagrebu Filozofski fakultet (Univ. of Zagreb, Faculty of Hum. and Social Sciences)	FFZG	HR
28	Københavns Universitet (University of Copenhagen)	UCPH	DK
29	Tartu Ülikool (University of Tartu)	UTART	EE
30	Helsingin Yliopisto (University of Helsinki)	UHEL	FI
31	Centre National de la Recherche Scientifique	CNRS	FR
32	Nyelvtudományi Kutatóközpont (Research Institute for Linguistics)	NYTK	HU
33	Stofnun Árna Magnússonar í íslenskum fræðum SAM (Árni Magnússon Inst. for Icelandic Studies)	SAM	IS
34	Fondazione Bruno Kessler	FBK	IT
35	Latvijas Universitātes Matemātikas un Informātikas institūts (Institute of Mathematics and Computer Science, University of Latvia)	IMCS	LV
36	Lietuvių Kalbos Institutas (Institute of the Lithuanian Language)	LKI	LT
37	Luxembourg Institute of Science and Technology	LIST	LU
38	Università ta Malta (University of Malta)	UM	MT
39	Stichting Instituut voor de Nederlandse Taal (Dutch Language Institute)	INT	NL
40	Språkrådet (Language Council of Norway)	LCNOR	NO
41	Instytut Podstaw Informatyki Polskiej Akademii Nauk (Polish Academy of Sciences)	IPIPAN	PL
42	Universidade de Lisboa, Faculdade de Ciências (University of Lisbon, Faculty of Science)	FCULisbon	PT
43	Institutul de Cercetări Pentru Inteligență Artificială (Romanian Academy)	ICIA	RO
44	University of Cyprus, French and European Studies	UCY	CY
45	Jazykovedný ústav Ľudovíta Štúra Slovenskej akadémie vied (Slovak Academy of Sciences)	JULS	SK
46	Institut Jožef Stefan (Jozef Stefan Institute)	JSI	SI
47	Centro Nacional de Supercomputación (Barcelona Supercomputing Center)	BSC	ES
48	Kungliga Tekniska högskolan (Royal Institute of Technology)	KTH	SE
49	Universität Zürich (University of Zurich)	UZH	CH
50	University of Sheffield	USFD	UK
51	Universidad de Vigo (University of Vigo)	UVIGO	ES
52	Bangor University	BNGR	UK

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Goals of the EP/EC Workshop</b>	<b>2</b>
<b>3</b>	<b>Previous Workshops and Related Activities</b>	<b>3</b>
3.1	STOA Workshop “Language Equality in the Digital Age – Towards a Human Language Project” (2017) . . . . .	3
3.2	EP Language Equality Conference (2018) . . . . .	3
3.3	EC Study “Opportunities and Challenges of Artificial Intelligence Technologies for the Cultural and Creative Sectors” (2022) . . . . .	4
3.4	EP Presentation “Towards Digital Language Equality in Europe by 2030” (2022)	4
<b>4</b>	<b>Conclusions</b>	<b>4</b>

## List of Acronyms

AI	Artificial Intelligence
CCS	Cultural and Creative Sectors
CULT	European Parliament's Committee on Culture and Education
EC	European Commission
ELE	European Language Equality ( <i>this project</i> )
EP	European Parliament
HLP	Human Language Project
ITRE	European Parliament's Committee on Industry, Research and Energy
MEP	Member of the European Parliament
META	Multilingual Europe Technology Alliance
META-NET	EU Network of Excellence to foster META
STOA	Science and Technology Options Assessment
TMNCL	Intergroup for Traditional Minorities, National Communities and Languages

## Abstract

A number of activities revolving around dissemination and communication, particularly related to the European Union, are addressed in WP4. The goal of the EP/EC workshop, as foreseen in the original ELE project proposal and grant agreement, was to present the main ELE results, especially the strategic research and innovation agenda and roadmap, and to discuss its concrete and sustainable implementation. The workshop, currently planned as a half-day workshop organised by the Science and Technology Options Assessment (STOA) unit, is targeted to take place in the second half of 2022 under the umbrella of the ELE follow-up project, ELE 2. Previous events, such as the STOA workshop in 2017, and ongoing network events have strengthened the ties to key stakeholders in the European Institutions and provided valuable experience for organising the upcoming workshop. While the workshop is targeted to take place later in 2022 it should also be noted that there has been continuous engagement with the EP/EC throughout the ELE project.

## 1 Introduction

WP4 contains all activities related to communication, dissemination and awareness raising on the political level both in the European Institutions and in the Member States. Its main objective is to help the topic of digital language equality gain visibility and ensure sustainability for the project goals and developed strategic agenda in Europe. This is achieved by (1) following an effective strategy and plan to support all communication and dissemination activities in ELE, (2) helping to organise open consultations and dialogues (together with WP2 and WP5), (3) leveraging existing structures to reach related stakeholder groups, (4) organising one large conference, and smaller outreach events, and (5) producing a final book publication with the main results of the ELE project. In addition to the large conference, META-FORUM 2022<sup>1</sup>, that took place on 8/9 June 2022, the consortium planned and organised a number of other smaller outreach events.

Deviating from the original plan, which had foreseen the EP/EC workshop to take place within ELE's project runtime, the workshop is targeted to take place in the second half of

<sup>1</sup> <https://www.european-language-grid.eu/meta-forum-2022/>

2022 under the umbrella of ELE 2. This document outlines the steps and initiatives taken to initiate the workshop. We highlight important conclusions from previous related workshops and presentations that serve as valuable input for the upcoming workshop, which is foreseen to be organised as a half day event organised by the Science and Technology Options Assessment (STOA) unit.

## 2 Goals of the EP/EC Workshop

The European Institutions have emphasised the need to strengthen multilingualism through technological innovations (European Parliament, 2018). Based on the Language Technology community's achievements of the last decade, the ELE consortium is keen on initiating discussions with the European Parliament and the European Commission on how to address the results and strategic recommendations of the ELE project. The goal we associate with the planned workshop is to present the main ELE results, especially the strategic agenda and roadmap and to discuss its concrete and sustainable implementation. A number of reliable and sustainable contacts with both European Institutions have been established since the beginning of the META-NET initiative in 2010.

When drafting the ELE timeline in June/July 2020, the ELE consortium initially targeted the workshop to take place in month 15 of the project (April 2022). However, various circumstances did not allow us to adhere to this timeline. In early 2022, many of the still existing COVID-19 restrictions made it extremely difficult to plan in-person events, which is still the preferred mode of the consortium for this kind of event. Face-to-face interaction is essential to develop strategic partnerships and network on a deeper level. Even though the uncertainty around COVID-19 remains, especially now with case numbers that are rising once again, we hope to run this event non-virtual in the second half of 2022.

While the language technology topic continues to be important for the European Parliament, the war in Ukraine, including the EU's responses, currently overshadows nearly all other discussions. Additionally, more than one year after Brexit, there has been a restructuring within the wider community. Longstanding support from allies has undergone changes.<sup>2</sup>

The consortium has continued its networking activities and cultivated the relationship with different groups within the European Institutions. Of particular importance has been the work with the Parliamentary Committee on Culture and Education (CULT).<sup>3</sup> One of CULT's main responsibilities lies in the protection and promotion of cultural and linguistic diversity. There have also been links to activities of the Committee on Industry, Research and Energy (ITRE).<sup>4</sup> And, the consortium has established ties with the Intergroup for Traditional Minorities, National Communities and Languages (TMNCL)<sup>5</sup> that serves as a meeting point for political groups, European institutions, non-governmental organisations and minority representatives.

<sup>2</sup> We would like to express our gratitude to Jill Evans who served as a Member of the European Parliament (MEP) for Wales from 1999 to 2020. She was the first person to use the Welsh language in debate at the European Parliament and has been vocal in supporting the struggles of minority languages and the importance of multilingualism. Jill Evans led the work that resulted in the 2018 EP resolution (European Parliament, 2018).

<sup>3</sup> <https://www.europarl.europa.eu/committees/en/cult/home/>

<sup>4</sup> <https://www.europarl.europa.eu/committees/en/itre/home/highlights>

<sup>5</sup> <https://www.europarl.europa.eu/about-parliament/en/organisation-and-rules/organisation/intergroups>

### 3 Previous Workshops and Related Activities

Based on previous experience, the preferred format for the workshop is a half-day STOA workshop.<sup>6</sup> The Science and Technology Options Assessment (STOA)<sup>7</sup> is a Panel of Members of the European Parliament (MEPs) with a mission to provide independent expert assessment of scientific and technological developments and policy options for informed political decision-making in the EP. STOA projects provide state-of-play overviews of different topics with scientific or technological dimension and assess a range of policy options for responsibly deploying while also promoting the use of existing and emerging technologies.

In the following, we briefly describe four recent activities that demonstrate our experience and show how we have paved the way for the actual organisation of the EP/EC workshop, targeted to take place in the second half of 2022.

#### 3.1 STOA Workshop “Language Equality in the Digital Age – Towards a Human Language Project” (2017)

On 10 January 2017, the STOA workshop “Language Equality in the Digital Age – Towards a Human Language Project” took place in the European Parliament (Brussels). The workshop was informed and made possible through META-NET’s finding that at least 21 European languages were in danger of digital extinction (Rehm and Uszkoreit, 2012). The title of the event hinted at a possible next step, which was also reflected and suggested in as well as supported by several of the presentations, i. e., to set up, in collaboration with the EC, EP and Member States, a large-scale, long-term initiative to carry out research, development and innovation activities with substantial funding towards a common strategic goal. Under the umbrella of what was then termed the “Human Language Project” (HLP), new breakthroughs were to be made in order to address the threat of digital language extinction, but also to provide solutions for European citizens, industry and administrations.

In March 2017, a corresponding study was published, commissioned by the EP (STOA, 2017), which suggested 11 policy recommendations, among others, to “reinforce the role of LT within the institutional framework of multilingualism-related bodies”, to “refocus and strengthen research in the Human Language Project” and to “bridge the technology gap between European languages”. The joint CULT/ITRE report (European Parliament, 2018), adopted by the EP on 11 Sept. 2018, is based on the results of the STOA study.

#### 3.2 EP Language Equality Conference (2018)

On 27 September 2018, a conference on Language Technologies and Digital Equality in a Multilingual Europe was held in the European Parliament in Brussels to discuss the equality of Europe’s languages in the digital age.<sup>8</sup> The conference gathered various presentations and software demonstrations that shed light on individual aspects of technology support for our multilingual continent, identified barriers, outlined the state of the art and showed perspectives for future developments. The conference was organised and chaired by MEP Jill Evans (Wales, UK).

Two weeks prior, in a plenary meeting on 11 September 2018, the European Parliament adopted the joint ITRE/CULT report “Language Equality in the Digital Age” with a majority of 592 votes in favour, 45 against and 44 abstentions. The adoption of the resolution with a

<sup>6</sup> The official application for the STOA workshop was submitted in early 2022 through a collaborator in the EP. The consortium is currently waiting for confirmation.

<sup>7</sup> <https://www.europarl.europa.eu/stoa/en/about/stoa-network>

<sup>8</sup> <https://www.greens-efa.eu/en/article/event/language-equality-in-the-digital-age>

“landslide” majority demonstrates the importance and relevance of the topic. This resolution is the most important political document for European language technology research in recent years and remains a great success of our joint efforts.

In the closing lecture of the event, “European Language Technologies – Past, Present and Future”, Georg Rehm summarised the alarming results of the META-NET study (which are still valid and will be further outlined in the upcoming ELE book). He also advocated that the EU Parliament, the European Commission and the Member States launch a large and long-term initiative to ensure new breakthroughs in basic research, applied research and also in the commercialisation of new research results.

### 3.3 EC Study “Opportunities and Challenges of Artificial Intelligence Technologies for the Cultural and Creative Sectors” (2022)

The study *Opportunities and Challenges of Artificial Intelligence Technologies for the Cultural and Creative Sectors*, commissioned by the European Commission and eventually published on February 2022, aims to identify opportunities in Artificial Intelligence (AI) for the Cultural and Creative Sectors (CCS), also taking into consideration the current challenges they pose, and suggests actions for CCS stakeholders as well as national and international policymakers in applying AI and promoting a culturally diverse offer of European works. Ten cultural and creative sectors were analysed including architecture, book publishing, fashion, film, museums, music, news media, performing arts, visual art and video games.

In Section 5.3 (starting from page 43), the report elaborates on opportunities and risks to language diversity. The report strongly emphasises that there is a need for Language Technologies “made in Europe, for Europe” (European Commission et al., 2022) and references the European Language Equality and European Language Grid projects as important relevant activities.

### 3.4 EP Presentation “Towards Digital Language Equality in Europe by 2030” (2022)

On 5 May 2022, Georg Rehm, coordinator of ELG and co-coordinator of ELE, presented the consortium’s vision of Digital Language Equality and how to achieve it to the Intergroup for Traditional Minorities, National Communities and Languages. The talk focused on the current state of both projects, ELG and ELE. It touched upon national AI strategies, the definition and metric of Digital Language Equality and results from the language reports. The findings and results were met with great interest. The full slide deck can be found in the appendix.

In this context, a recent report from the Council of the European Union Education, Youth, Culture and Sport Council<sup>9</sup>, published on 4-5 April 2022 was shared and discussed. The experts ask to reinforce intercultural exchanges and develop a European strategy for cultural and creative industries. The conclusions also call for the establishment of an ambitious digital policy for the development of language technologies. The EU wants to take advantage of new technologies to foster multilingualism, which nurtures cultural exchanges and facilitates access to culture.

## 4 Conclusions

The already existing close collaboration with the Parliamentary Committee on Culture and Education (CULT), the Committee on Industry, Research and Energy (ITRE) and the Inter-

<sup>9</sup> <https://www.consilium.europa.eu/en/meetings/eycs/2022/04/04-05/>



group for Traditional Minorities, National Communities and Languages (TMNCL) have laid important groundwork for the EP/EC workshop, which is planned to take place in the second half of 2022. The experience from previous workshops and collaborations has been very helpful in helping to draft a concept for the upcoming workshop. In 2017, the results of the STOA workshop “Language Equality in the Digital Age – Towards a Human Language Project” inspired the corresponding STOA study that was published a year later. The digital language equality topic gained momentum and visibility afterwards. We expect the planned STOA workshop later this year to initiate similar substantial impact.

## References

- European Commission, Directorate-General for Communications Networks, Content and Technology, K. Izsak, A. Terrier, S. Kreutzer, T. Strähle, C. Roche, M. Moretto, S. Sorensen, M. Hartung, K. Knaving, M. Johansson, M. Ericsson, and D. Tomchak. *Opportunities and challenges of artificial intelligence technologies for the cultural and creative sectors*. 2022. doi: doi/10.2759/144212.
- European Parliament. Language Equality in the Digital Age. European Parliament resolution of 11 September 2018 on Language Equality in the Digital Age (2018/2028(INI). [http://www.europarl.europa.eu/doceo/document/TA-8-2018-0332\\_EN.pdf](http://www.europarl.europa.eu/doceo/document/TA-8-2018-0332_EN.pdf), 2018.
- Georg Rehm and Hans Uszkoreit, editors. *META-NET White Paper Series: Europe's Languages in the Digital Age*, 32 volumes on 31 European languages, Heidelberg etc., 2012. Springer.
- STOA. Language equality in the digital age – Towards a Human Language Project. STOA study (PE 598.621), IP/G/STOA/FWC/2013-001/Lot4/C2, March 2017. Carried out by Iclaves SL (Spain) at the request of the Science and Technology Options Assessment (STOA) Panel, managed by the Scientific Foresight Unit (STOA), within the Directorate-General for Parliamentary Research Services (DG EPRS) of the European Parliament, March 2017. <http://www.europarl.europa.eu/stoa/>.

## Appendix



# EUROPEAN LANGUAGE EQUALITY



## EUROPEAN LANGUAGE GRID


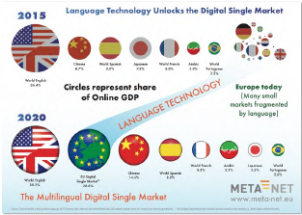
## Towards Digital Language Equality in Europe by 2030

Prof. Dr. Georg Rehm (DFKI GmbH, Germany) – georg.rehm@dfki.de  
Coordinator ELG, Co-coordinator ELE

05-05-2022 European Parliament – Intergroup for Traditional Minorities, National Communities and Languages  
<http://www.european-language-grid.eu>  
<http://www.european-language-equality.eu>

## Point of Departure: Multilingualism in Europe

- Multilingualism is at the heart of the European idea
- 24 official EU languages – they all have the same status
- Dozens of co-official, regional and minority languages as well as languages of immigrants and trade partners
- Many economic, social and technical challenges, e.g.,
  - The Digital Single Market needs to be multilingual
  - Cross-border, cross-lingual, cross-cultural communication

2



The timeline illustrates the progression of the European Language Grid (ELG) and its associated initiatives. Key milestones include the founding of META-NET in 2020, the launch of the Strategic Research and Innovation Agenda (SRIA) in 2021, and the ongoing efforts to achieve digital language equality by 2030. The timeline also highlights the role of the European Language Grid (ELG) consortium and the European Language Equality (ELE) initiative.

3



### META-NET Language White Papers “Europe’s Languages in the Digital Age”

2012

Basque	Galician	Polish*
Bulgarian*	German*	Portuguese*
Catalan	Greek*	Romanian*
Croatian*	Hungarian*	Serbian
Czech*	Icelandic	Slovak*
Danish*	Irish*	Slovene*
Dutch*	Italian*	Spanish*
English*	Latvian*	Swedish*
Estonian*	Lithuanian*	Welsh
Finnish*	Maltese*	
French*	Norwegian	


\* Official EU language

4



	Excellent support	Good support	Moderate support	Fragmentary support	Weak or no support
<b>Machine Translation</b>		English	French, Spanish	Catalan, Dutch, German, Hungarian, Italian, Polish, Romanian	Basque, Bulgarian, Croatian, Czech, Danish, Estonian, Finnish, Galician, Greek, Icelandic, Irish, Latvian, Lithuanian, Maltese, Norwegian, Portuguese, Serbian, Slovak, Slovene, Swedish, Welsh
<b>Text Analytics</b>		English	Dutch, French, German, Italian, Spanish	Basque, Bulgarian, Catalan, Czech, Danish, Finnish, Galician, Greek, Hungarian, Norwegian, Polish, Portuguese, Romanian, Slovak, Slovene, Swedish	Croatian, Estonian, Icelandic, Irish, Latvian, Lithuanian, Maltese, Serbian, Welsh
<b>Speech</b>		English	Czech, Dutch, Finnish, French, German, Italian, Portuguese, Spanish	Basque, Bulgarian, Catalan, Danish, Estonian, Galician, Greek, Hungarian, Irish, Norwegian, Polish, Serbian, Slovak, Slovene, Swedish	Croatian, Icelandic, Latvian, Lithuanian, Maltese, Romanian, Welsh
<b>Resources</b>		English	Czech, Dutch, French, German, Hungarian, Italian, Polish, Spanish, Swedish	Basque, Bulgarian, Catalan, Croatian, Danish, Estonian, Finnish, Galician, Greek, Norwegian, Portuguese, Romanian, Serbian, Slovak, Slovene	Icelandic, Irish, Latvian, Lithuanian, Maltese, Welsh

5



2012

Level of support

Excellent

Good

Moderate

Fragmentary

Weak/none

English, French, Spanish, Dutch, German, Italian, Czech, Hungarian, Polish, Finnish, Swedish, Portuguese, Basque, Bulgarian, Danish, Galician, Greek, Norwegian, Romanian, Slovak, Slovene, Estonian, Serbian, Croatian, Irish, Icelandic, Latvian, Lithuanian, Maltese, Welsh

Languages with names in red have little or no MT support

META-NET

Source: META-NET White Paper Series: Europe’s Languages in the Digital Age. Springer, Heidelberg, New York, Dordrecht, London, September 2012. Georg Behn and Hans Uszkoreit (series editors)

6

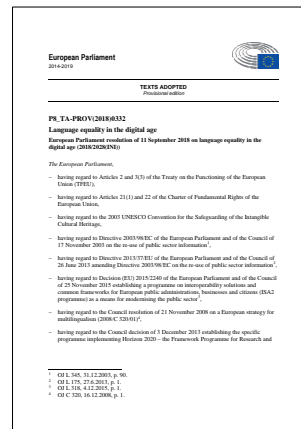
## European Parliament Resolution (2018)

EP Resolution “Language equality in the digital age”  
P8\_TA(2018)0332 – partially based on the STOA study

Voting (11 Sept. 2018): **592 yes** – 45 no

**Recommendations addressed by European Language Grid:**

- 27. Europe has to secure its leadership in language-centric AI**
- 29. Create a European LT platform for sharing of services**
- 41. Enable and empower European SMEs to use LTs**
- 26. ICT integrators should be given economic incentives for LT**
- 32. Set up LT financing platform; emphasise R&D in Deep NLU**
- 40. Develop investment instruments and accelerator programs**



ELG European Language Grid

7



**EUROPEAN  
LANGUAGE  
GRID**



**EUROPEAN  
LANGUAGE  
EQUALITY**

### Development of a cloud-based technology platform

*Horizon 2020 project: January 2019 – June 2022 (ELG)*

#### Current state of play:

- Platform has been developed and populated with 12,000+ tools, services and resources
- Legal entity to be established for further maintenance and extension of the platform and promotion towards the community
- This talk: brief overview

### Development of a strategic research, innovation and implementation agenda and roadmap towards digital language equality in Europe by 2030

*EU/EC pilot project: January 2021 – June 2022 (ELE)*

*EU/EC pilot project: July 2022 – June 2023 (ELE2)*

#### Current state of play:

- Produced a large number of reports and results
- These need to be put together and assembled into the strategic agenda (also to be published as a book)
- This talk: overview of selected preliminary results

ELG European Language Grid

8



**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



ELG European Language Grid

9

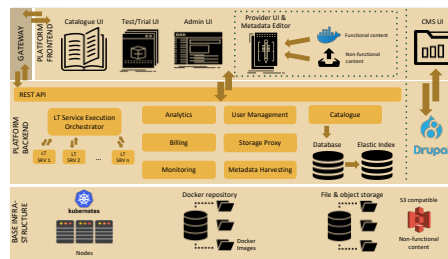
## Current State of Play (May 2022)

### Release 2 (May 2022): 12,000+ Resources

- 6267 corpora and data sets
  - To train machine learning models, to perform linguistic analyses etc.
- 3099 services and tools
  - 770 functional services directly within ELG: Machine translation, automated speech recognition, summarisation, information extraction, text-to-speech synthesis etc.
- 2238 lexical/conceptual resources
- 427 lang. descriptions, grammars, models
- 1778 organisations (research orgs., companies)

*These resources are provided by members of the European Language Technology community.*

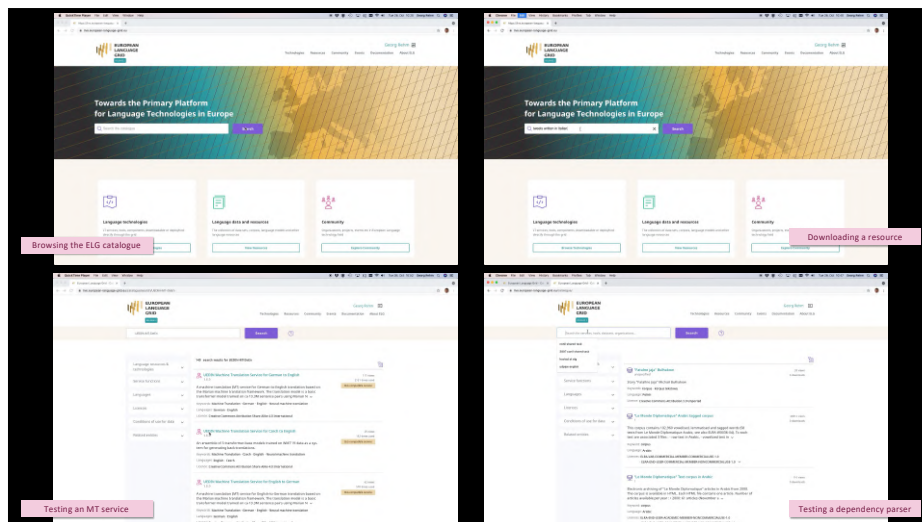
ELG European Language Grid



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

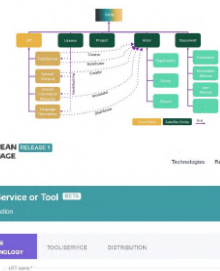
EUROPEAN  
LANGUAGE  
GRID

10



## European Language Grid – What can you do with it?

- **Data consumers** can search and browse the ELG catalogue
  - for different types of data, language processing services, projects and organisations
  - download data (depending on access conditions)
- **Service consumers** can try out and test language processing services
  - call a service from the command line, integrate it into workflow
  - view code samples
  - current API support: MT, IE, ASR, TTS, text classification
- **Data and service consumers** can use a Python-based API for accessing the ELG catalogue, searching and directly fetching datasets to feed them into, e.g., their model training pipeline
- **Data providers** can upload resources and register metadata descriptions using a semantic schema
  - a dashboard provides an overview of own resources including their status (*draft* → *published*)
- **Language Technology providers** can provide their service/tool as a Docker container
  - Different options: (a) Docker image in ELG, can be run from ELG; (b) Docker image in ELG as proxy talks to remote service



ELG European Language Grid

12



## META-FORUM Conference Series

**META-FORUM 2022** – June 08-10, Brussels, Belgium  
**Joining the European Language Grid**

**META-FORUM 2021** – November 15-17, *virtual conference*  
**Using the European Language Grid**

**META-FORUM 2020** – December 01-03, *virtual conference*  
**Piloting the European Language Grid**

**META-FORUM 2019** – October 08/09, Brussels, Belgium  
**Introducing the European Language Grid**

**META-FORUM 2017** – November 13/14, Brussels, Belgium  
**Towards a Human Language Project**

**META-FORUM 2016** – July 04/05, Lisbon, Portugal  
**Beyond Multilingual Europe**

**META-FORUM 2015** – April 27, Riga, Latvia  
**Technologies for the Multilingual Digital Single Market**

**META-FORUM 2013** – September 19/20, Berlin, Germany  
**Connecting Europe for New Horizons**

**META-FORUM 2012** – June 20/21, Brussels, Belgium  
**A Strategy for Multilingual Europe**

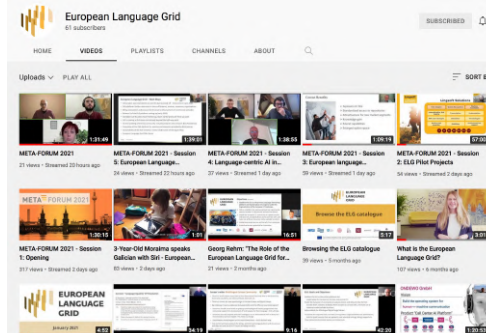
**META-FORUM 2011** – June 27/28, Budapest, Hungary  
**Solutions for Multilingual Europe**

**META-FORUM 2010** – November 17/18, Brussels, Belgium  
**Challenges for Multilingual Europe**

ELG European Language Grid

All META-FORUM 2020 and META-FORUM 2021 sessions are available on the European Language Grid YouTube channel:

<https://www.youtube.com/channel/UCarEHmsWT2JslcvWkblL4A>



13

## European Language Equality (ELE)



ELG European Language Equality

**Consortium:** 52 partners from all over Europe

**Coordinator:** ADAPT Centre (Dublin City University)

**Co-Coordinator:** DFKI

**Objective:** *development of a strategic research, innovation and implementation agenda to achieve digital language equality in Europe by 2030*

**Started on 1 January 2021**

**Runtime:** 18 months – ELE and ELG will both finish up in June 2022 with a joint META-FORUM 2022

ELE2 will start in July 2022 and run until June 2023

<http://www.european-language-equality.eu>

14

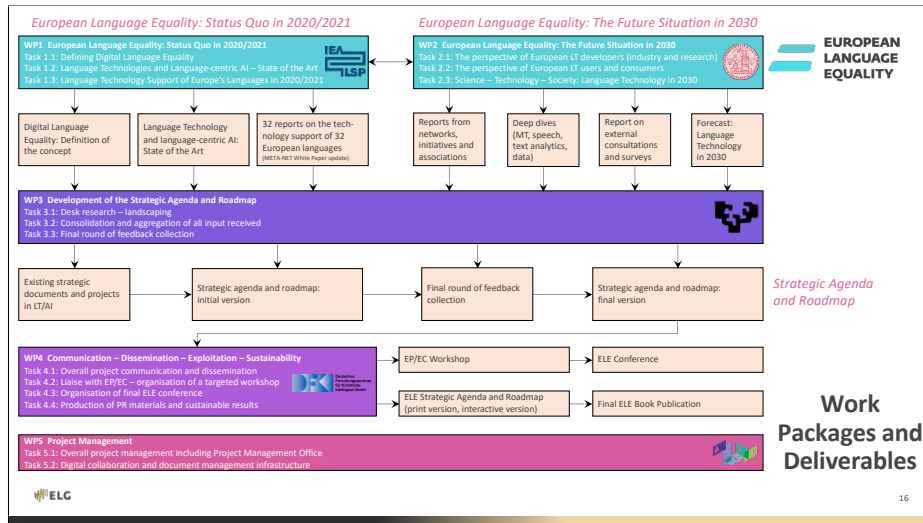
## Approach



- **Main result of the project:** **Strategic Research, Innovation and Implementation Agenda and a Roadmap for Achieving full Digital Language Equality in Europe by 2030.**
- Research partners have produced updates of the 30+ META-NET White Papers.
- Various networks and pan-European initiatives have produced one report each in which they present their own position, needs, wishes, demands, visions etc.
- Various companies have produced four technical deep dives for the four main tech areas.
- Number of languages taken into account: 77
- Total number of project deliverables: 66

ELG European Language Equality

15

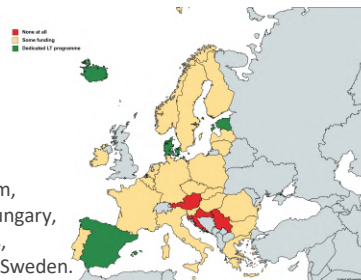


### Current State of Play in the Project (Selection)

- **Language Technology and National AI Strategies:** performed Europe-wide landscaping.
- **Digital Language Equality (DLE):** concept and metric defined, distinction between *technological factors* and *contextual factors*.
- **Dashboard – Dynamic DLE Computation and Visualisation:** based on the 12,000+ metadata records in ELG we can compute the DLE metric for all European languages and visualise the results. We can even monitor the development of DLE in Europe over time.
- **ELE Language Reports – Updates of the META-NET White Papers:** qualitative assessment of the gaps and needs of the various European languages.

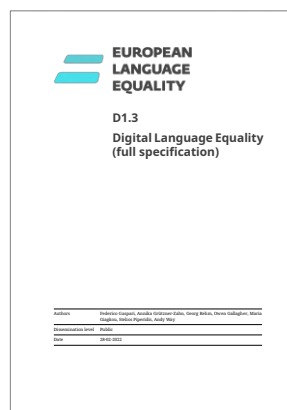
### Language Technology and National AI Strategies

- All 30 countries surveyed have a national AI strategy.
- Funding for Language Technology through AI only available in six countries: Belgium, Denmark, Estonia, France, Germany and Malta.
- No Language Technology-related funding at all in Austria, Croatia, Serbia.
- Limited Language Technology-related funding in Belgium, Bulgaria, Czechia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Lux., Malta, Netherlands, Norway, Poland, Portugal, Romania, Slovakia, Slovenia, Sweden.
- Dedicated LT programmes in Denmark, Estonia, Iceland and Spain.
- There is an immense need for pan-European coordination and for putting LT into the national AI strategies and into the policies for language, cultural and technological development



### Digital Language Equality Metric

- Digital Language Equality Metric
  - Provide theoretical basis to achieve DLE in Europe by 2030
  - Key ingredient of the evidence-based SRIA and roadmap
  - Enable comparisons among Europe's languages
  - Identify gaps
- Well-defined quantifiers, measures and indicators
- Compatibility and interoperability with ELG metadata schema, i.e., the DLE definition is fully aligned with ELG
- Intended to guide and prioritise future LT development and LR creation, collection, curation



### Digital Language Equality: Definition



**Digital Language Equality (DLE)** is the state of affairs in which all languages have the technological support and situational context necessary for them to continue to exist and to prosper as living languages in the digital age.

### Digital Language Equality Metric

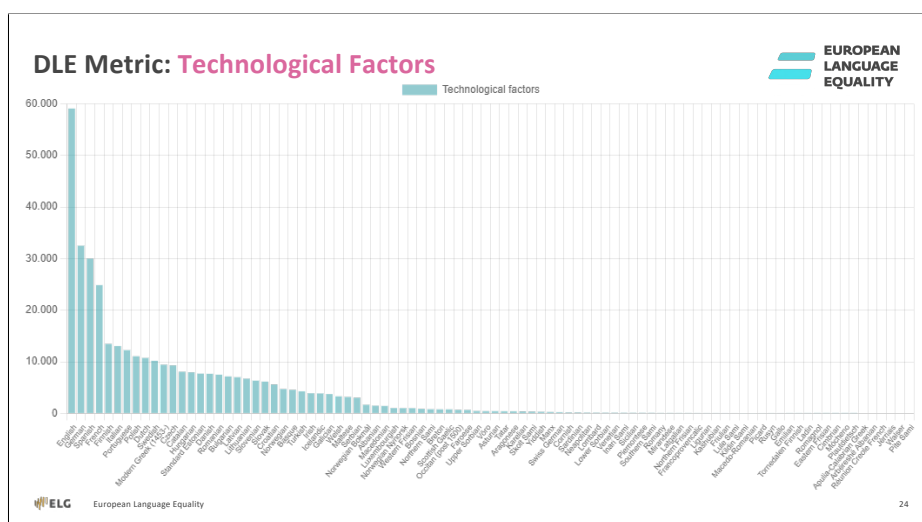
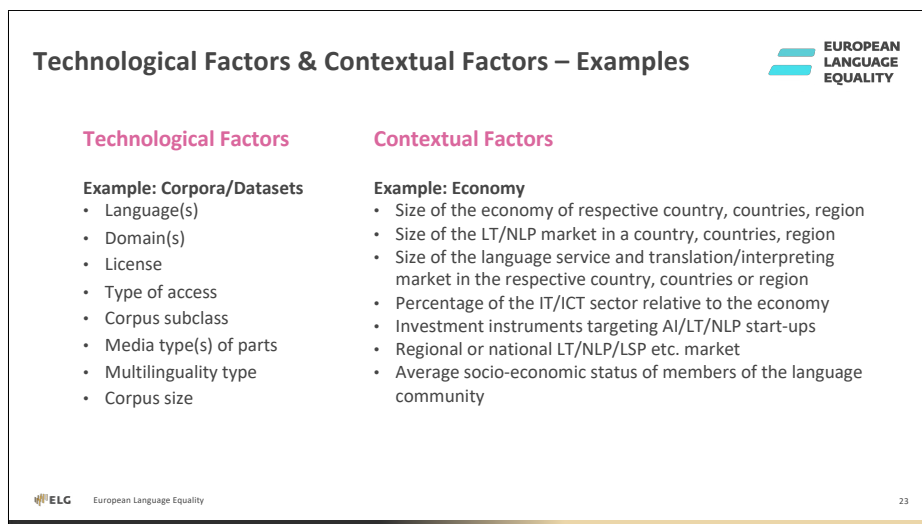
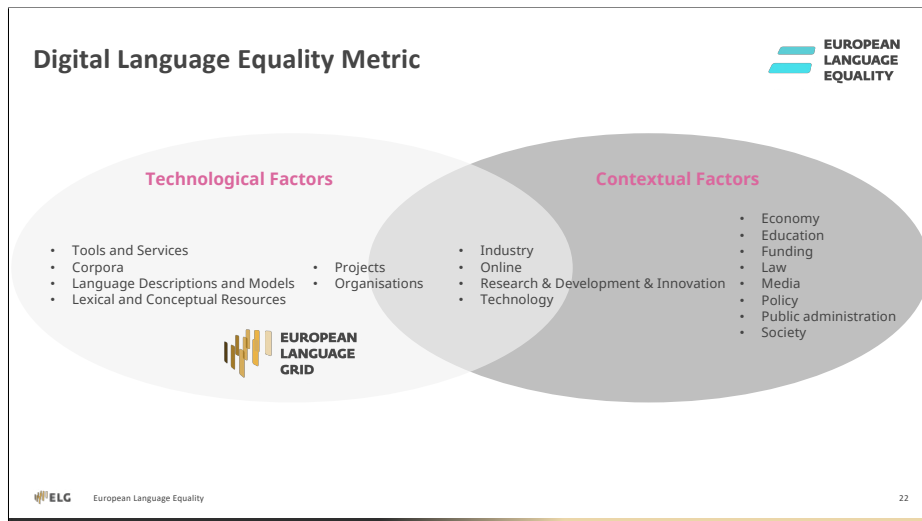


The DLE metric is a measure that reflects the digital readiness of a language and its contribution to the state of technology-enabled multilingualism, tracking its progress towards the goal of DLE.

The DLE metric is computed for each language on the basis of *various factors*, grouped into **technological factors** (e.g., the available language resources, tools and technologies) and a range of situational **contextual factors** (e.g., societal, economic, educational, industrial).

*dynamic nature, score can be (regularly) updated*





## DLE Metric: Contextual Factors – Final Set of Indicators



Class	Factor(s)	Indicators
Economy	Size of the economy	Gross domestic product (GDP)
	Size of the ICT sector	Percentage of the ICT sector in GDP and ICT service exports
Education	Students in LT/language courses of study	Number of students
	Inclusion in education	Percentage of population from a foreign country with tertiary education
Industry	Companies developing LTs	Number of companies in the ICT sector
Law	Legal status and legal protection	Scores extracted out of the legal status
Online	Wikipedia pages	Number of Wikipedia pages
Research & Development & Innovation	Innovation capacity	Innovation Index
	Number of papers	Number of entries in Glottolog
Society	Size of language community	Number of speakers
	Usage of social media	Number of individuals & percentage of speakers using social media
Technology	Digital connectivity and internet access	Percentage of households with broadband access

Weighting of the results via the two factors: *Number of Speakers and Legal Status*

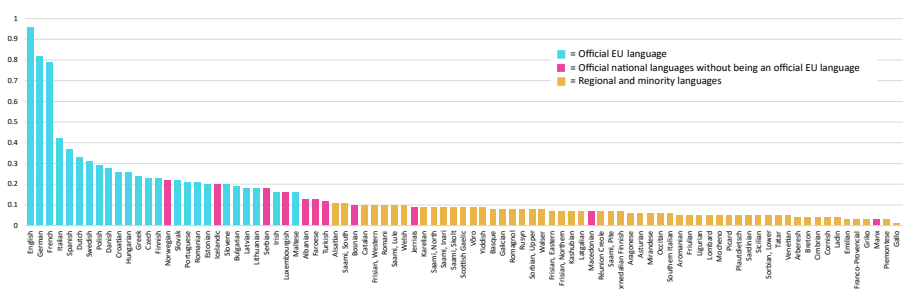
Booster factor: *Member of the European Economic Area + Penalty factor: Vitality state and Competition with English as another official national language*



Announcing the ELG Open Calls: Open Calls for Pilot Projects

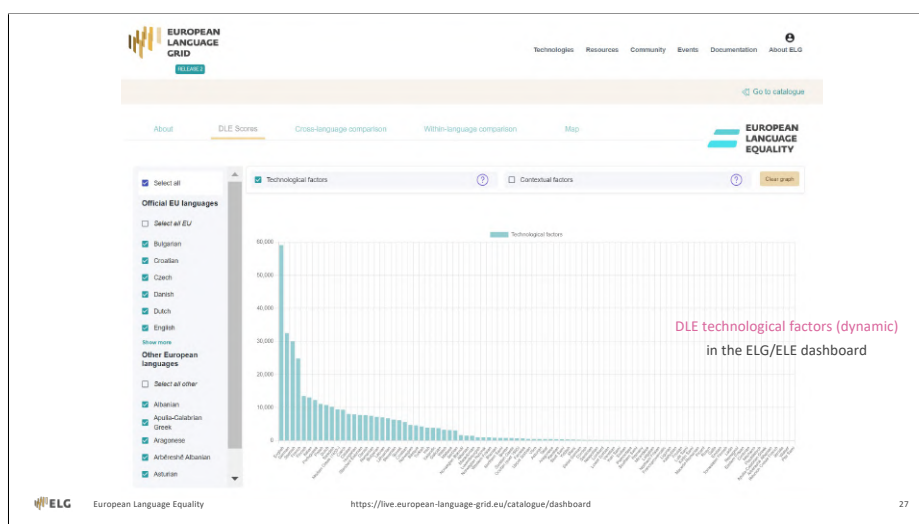
25

## DLE Metric: Contextual Scores

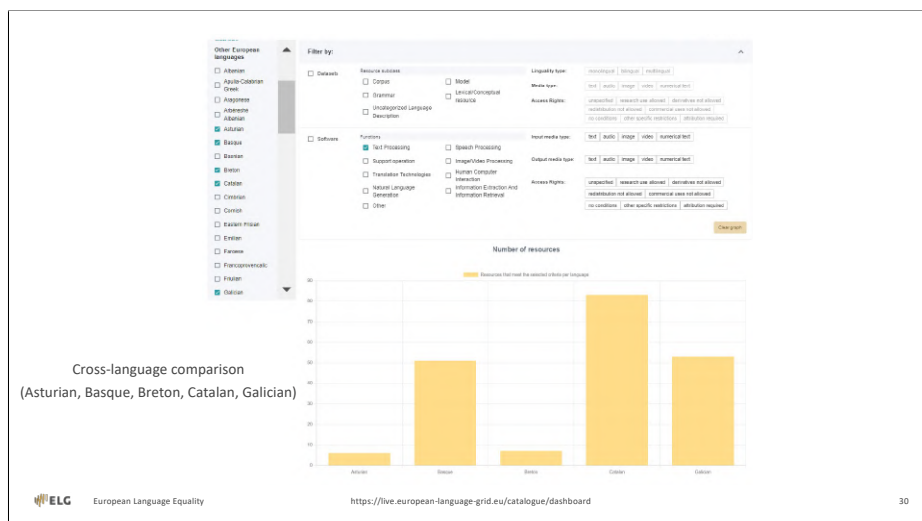
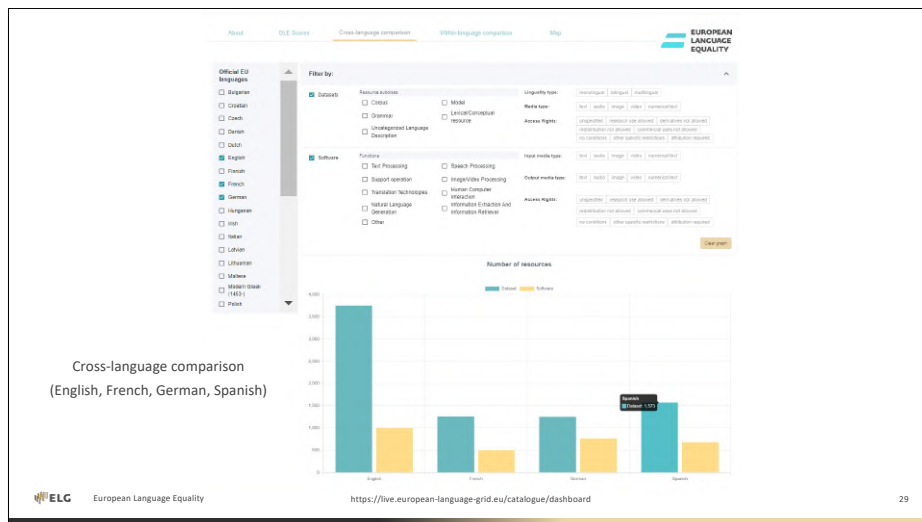
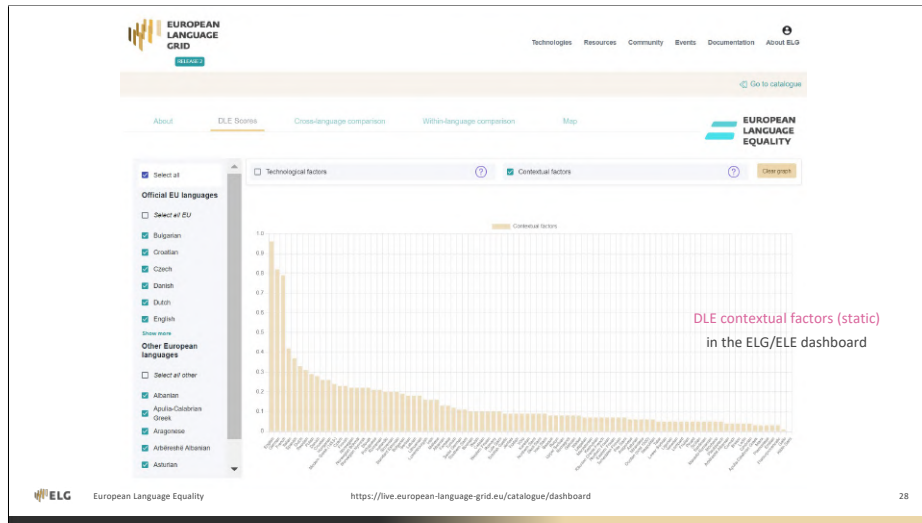


Announcing the ELG Open Calls: Open Calls for Pilot Projects

26



27

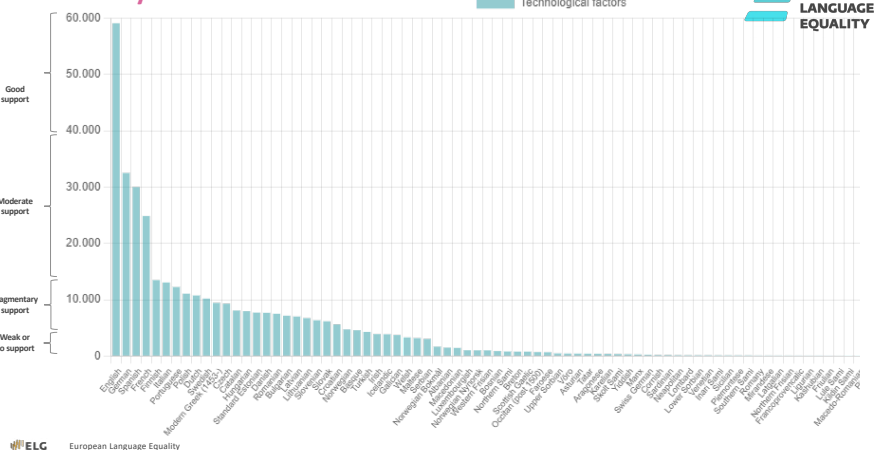


## Language Reports

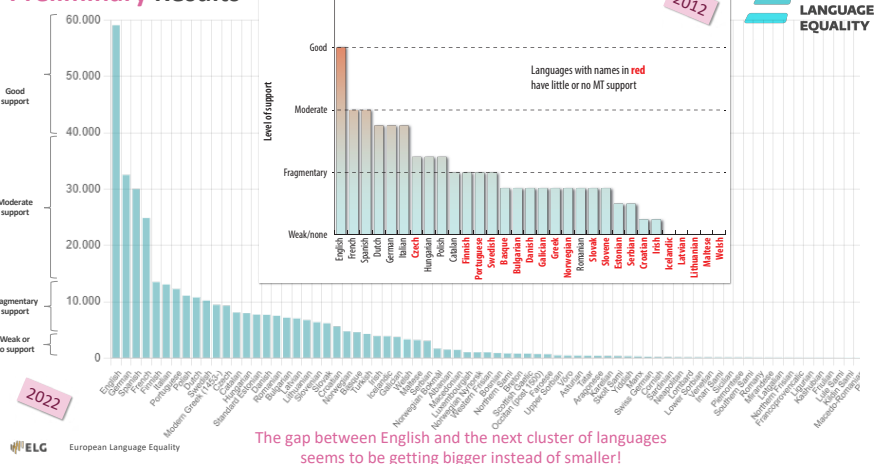


- In the tradition of the META-NET White Paper Series (published in 2012)
- Quantitative and qualitative evaluations of the level of support each language receives through technology:
  - Brief introduction to the LT field
  - General facts about the language
  - Availability of LRTs based on the combined collection of ELG and ELE resources
  - Support it receives through dedicated funding programs and projects
  - Participation in research infrastructures
- All reports available at <https://european-language-equality.eu/deliverables/>
- Stand-alone digital publications; parts will be included in the final ELE agenda and book, which will be published in the second half of 2022.

## Preliminary Results



## Preliminary Results



## Next Steps and Recommendations

preliminary

- 2012: At least 21 European languages in danger of **digital extinction** – the real number is *much* higher.
- Europe wants AI to be “based on and respecting European values”. Let’s take this seriously and develop AI-based technologies for *all* our languages, given that they’re all on the same level. Right now, **European Language Technology development and support is not equal, i.e., not based on European values.**
- We advocate for a **shared, long-term European funding programme** – between the EC and member states. All languages, i.e., countries and regions need to be brought on board (no more sporadic, partial support).
- Extremely strong need for **large language models (LLMs) for all European languages.**
- Prerequisites: **powerful data and compute infrastructure and vast amounts of FAIR, open language data** for all European languages, uses, apps. etc. – there *is* lots of data but it’s simply not available right now.
- **Global LT/NLP market size by 2025 is enormous (\$US 45B):** we want Europe to be a **key player.**
- European digital language and data processing infrastructure crucial for **digital sovereignty.**
- Need for pan-European coordination and a joint scientific goal: **Deep Natural Language Understanding.**

ELG European Language Grid

34

## There are also good news: More and more interest in this topic ...

### Council of the European Union

#### Education, Youth, Culture and Sport Council, 4-5 April 2022

#### Main Results

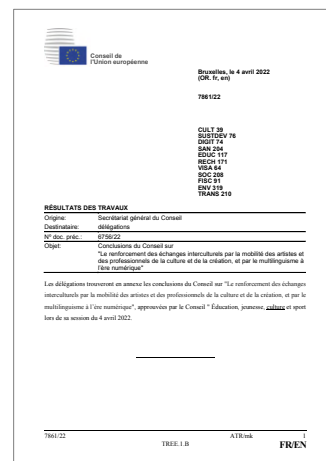
*Approval of Council conclusions on reinforcing intercultural exchanges and on a European strategy for cultural and creative industries*

**The conclusions also call for the development of an ambitious digital policy for the development of language technologies, [...].**  
**The EU wants to take advantage of new technologies to foster multilingualism, which nurtures cultural exchanges and facilitates access to culture.**

<https://www.consilium.europa.eu/en/meetings/eycs/2022/04/04-05/>

ELG European Language Grid

35



Thank you!



The European Language Grid has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 825627 (ELG).  
 The European Language Equality project has received funding from the European Union under grant agreement No LC-01641480 – 101018166 (ELE).

Prof. Dr. Georg Rehm (DFKI GmbH, Germany) – georg.rehm@dfki.de  
 Coordinator ELG, Co-coordinator ELE

05-05-2022 European Parliament – Intergroup for Traditional Minorities,  
 National Communities and Languages  
<http://www.european-language-grid.eu>  
<http://www.european-language-equality.eu>



With many thanks to all  
 colleagues in the ELG and  
 ELE teams and consortia!

