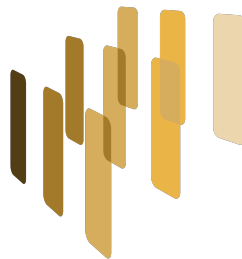




**EUROPEAN  
LANGUAGE  
EQUALITY**



**EUROPEAN  
LANGUAGE  
GRID**

## **European Language Equality and European Language Grid: An Overview**

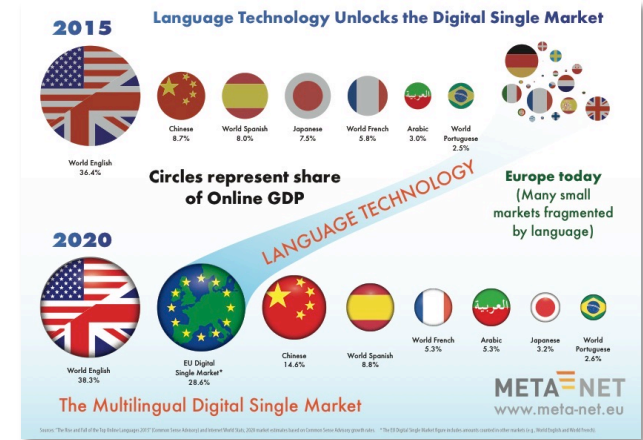
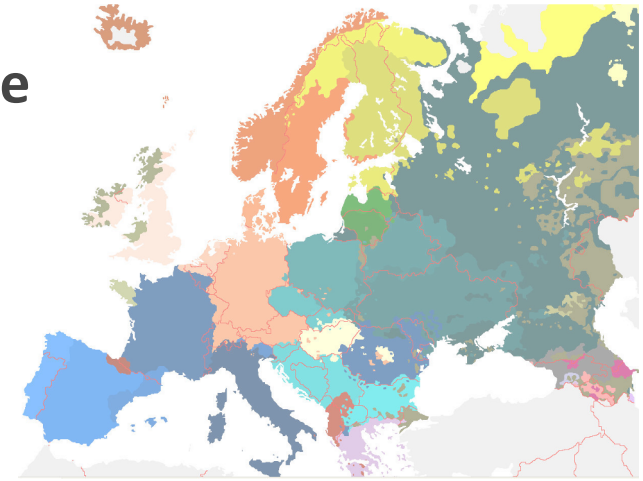
Katrin Marheinecke (Project Manager ELG, DFKI Berlin)

02-06-2021 Latvian Conference “Digital resources and tools for education”

<http://www.european-language-equality.eu> and <https://www.european-language-grid.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
  - The Digital Single Market needs to be multilingual
  - Cross-border, cross-lingual, cross-cultural communication
  - Fragmentation of the LT market and landscape
  - Most languages are digitally under-resourced

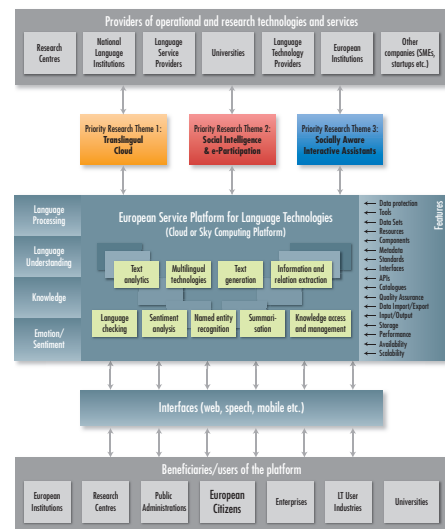


# Motivation and Context

- The European Language Technology landscape is very fragmented.
- The European LT community has been demanding a **European Language Technology platform** for several years.
- **META<sup>3</sup>NET** vision: **European Service Platform for Language Technologies**.
- First mention in the **Strategic Research Agenda for Multilingual Europe 2020** (published in early 2013).



META<sup>3</sup> FORUM 2013



The concept was later refined in the three LT SRIAs (2015, 2016, 2017).



META<sup>3</sup> FORUM 2015



META<sup>3</sup> FORUM 2016



META<sup>3</sup> FORUM 2017

# Context: “Language Equality” EP Resolution

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

## Selected recommendations addressed by ELE and ELG:

- 25. Establish a large-scale, long-term coordinated funding programme for research, development and innovation in the field of language technologies, at European, national and regional levels, tailored specifically to Europe’s needs and demands
- 27. Europe has to secure its leadership in language-centric AI
- 29. Create a European LT platform for sharing of services

European Parliament

2014-2019



TEXTS ADOPTED  
Provisional edition

**P8\_TA-PROV(2018)0332**

**Language equality in the digital age**

**European Parliament resolution of 11 September 2018 on language equality in the digital age (2018/2028(INI))**

*The European Parliament,*

- having regard to Articles 2 and 3(3) of the Treaty on the Functioning of the European Union (TFEU),
- having regard to Articles 21(1) and 22 of the Charter of Fundamental Rights of the European Union,
- having regard to the 2003 UNESCO Convention for the Safeguarding of the Intangible Cultural Heritage,
- having regard to Directive 2003/98/EC of the European Parliament and of the Council of 17 November 2003 on the re-use of public sector information<sup>1</sup>,
- having regard to Directive 2013/37/EU of the European Parliament and of the Council of 26 June 2013 amending Directive 2003/98/EC on the re-use of public sector information<sup>2</sup>,
- having regard to Decision (EU) 2015/2240 of the European Parliament and of the Council of 25 November 2015 establishing a programme on interoperability solutions and common frameworks for European public administrations, businesses and citizens (ISA2 programme) as a means for modernising the public sector<sup>3</sup>,
- having regard to the Council resolution of 21 November 2008 on a European strategy for multilingualism (2008/C 320/01)<sup>4</sup>,
- having regard to the Council decision of 3 December 2013 establishing the specific programme implementing Horizon 2020 – the Framework Programme for Research and

<sup>1</sup> OJ L 345, 31.12.2003, p. 90.

<sup>2</sup> OJ L 175, 27.6.2013, p. 1.

<sup>3</sup> OJ L 318, 4.12.2015, p. 1.

<sup>4</sup> OJ C 320, 16.12.2008, p. 1.



# European Language Grid (ELG)



**Objective:** establishing ELG as the primary platform in Europe for commercial and non-commercial, industry-related LTs (functional and non-functional)

**Consortium:** 9 partners from all over Europe

**Coordinator:** German Research Center for Artificial Intelligence (DFKI GmbH, Germany)

**Start** on 1 January 2019

**Runtime:** 42 months –  
(due to Covid-19-related extension)

# European Language Equality (ELE)



## EUROPEAN LANGUAGE EQUALITY

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

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

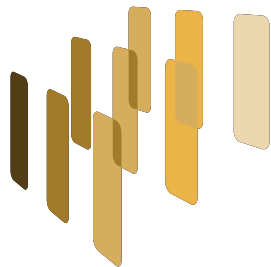
**Coordinator:** ADAPT Centre (Dublin City University)

**Co-coordinator:** DFKI GmbH

**Start on 1 January 2021**

**Runtime:** 18 months –

ELE and ELG will both end in June 2022



European Language Grid

The ELG Project



Kick-off meeting, 22/23 January 2019

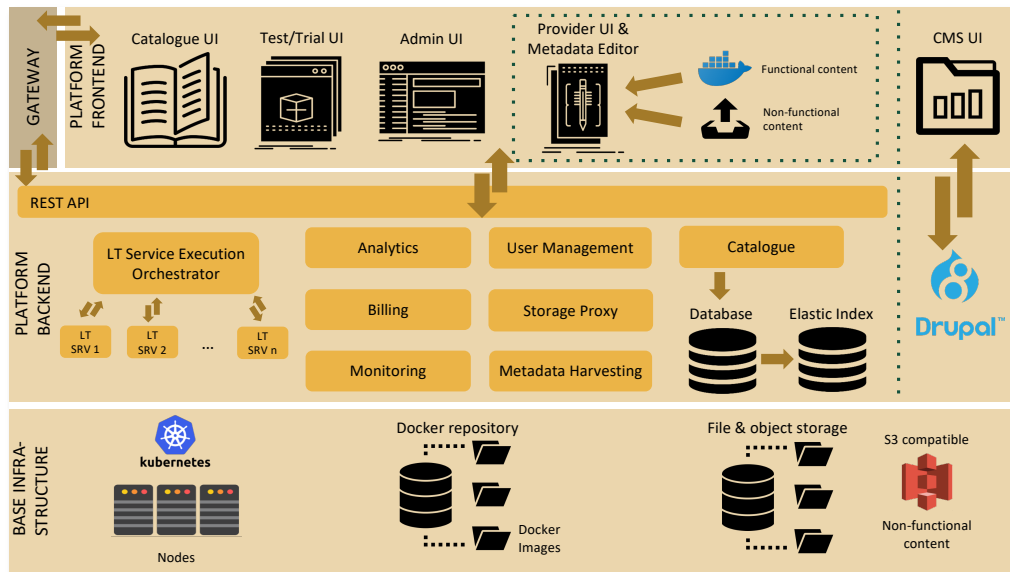
## Objectives (Selection)

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, both functional and non-functional.
3. Enable the European LT community to upload services and data sets into the ELG, 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.

# European Language Grid – Current State of Play (Mid 2021)

- User registration, authentication, authorisation
- User categories, respective rights and policies
- LT metadata upload and editing facilities
- Metadata conversion and harvesting
- LT service registration, integration
- LT service try out and execution
- LT data browsing, searching, upload, download
- Online documentation
- Current state of play in the European Language Grid in ELG Release 2:
  - 289 functional services and tools
  - 2576 corpora
  - 753 lexical/conceptual resources
  - 8 language descriptions

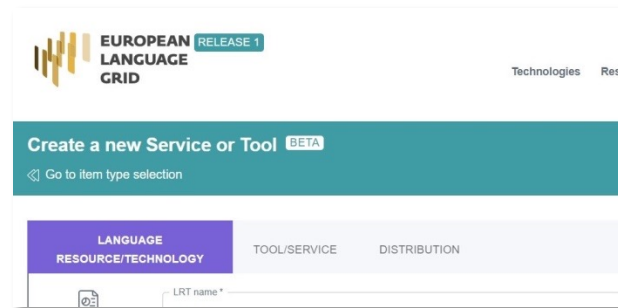
- *Facilities for consumers, esp. of data and services*
- *Facilities for providers, esp. of data and services*



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

# 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: Docker image directly in ELG; Docker image in ELG talks to remote service



# 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 2019) – Brussels, Belgium

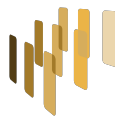
# Community: NCCs and LTC

## 32 National Competence Centres (NCCs)

- Strong international network of national networks to broaden ELG's reach, identify content for the ELG and interest companies in using the ELG.
- Main goal: *support the mission of the ELG project.*

## European LT Council (LTC)

- A pan-European body, in which strategic LT-related matters can be discussed and coordinated.
- Main goal: *represent and support European LT community.*



### Austria



DAGMAR GROMANN

Centre for Translation Studies, University of Vienna

### Belgium



WALTER DAELEMANS

Computational Linguistics and Psycholinguistics, University of Antwerp

### Bulgaria



SVETLA KOEVA

Institute for Bulgarian Language, Bulgarian Academy of Sciences

### Croatia



MARKO TADIĆ

Department of Linguistics, University of Zagreb

### Cyprus



DORA LOIZIDOU

University of Cyprus

### Czech Republic



JAN HAJIČ

Institute of Formal and Applied Linguistics, Charles University Prague

### Denmark



BOLETTE SANDFORD PEDERSEN

Centre for Language Technology, Department of Nordic Studies and Linguistics, University of Copenhagen

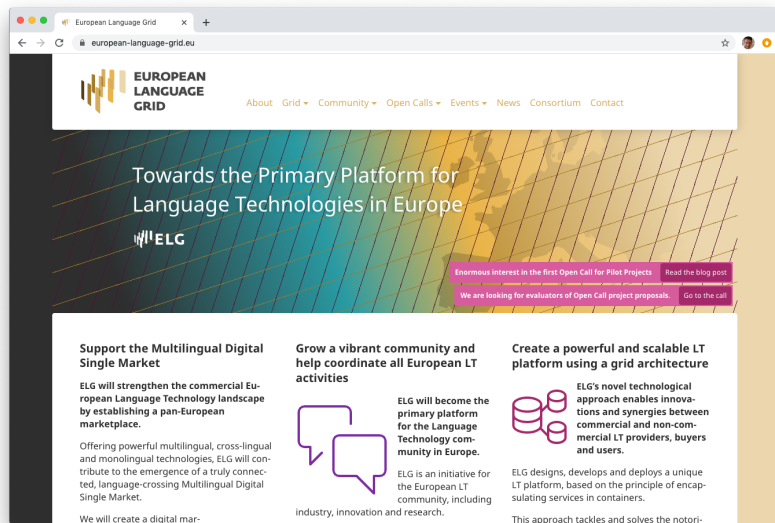


# European Language Grid: Sustainable Operational Model & Legal Entity

- ELG is supposed to be a long-term, sustainable initiative – a legal entity is needed.
- The technical and operational requirements – high availability and performance, SLAs, billing, support etc. – create non-trivial costs: hosting; bandwidth; ELG team; legal; etc.
- Several ways and approaches of covering the costs on a long-term basis have been identified
- Consensus for a sustainable operational model to be established in fall 2021
- First concepts are being prepared.
- Q4/2021: Establish legal entity (probably with a soft start)

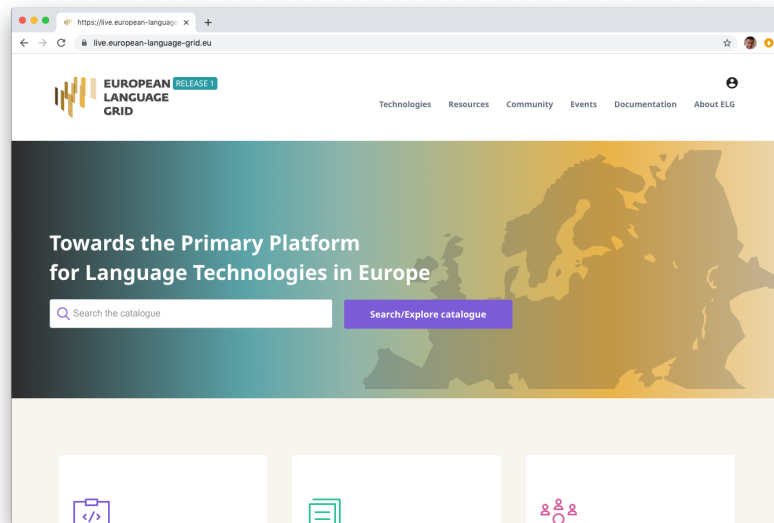


# European Language Grid



<https://www.european-language-grid.eu>

- ELG project website
- Typical EU project information plus a lot of additional content (events, open calls, community, NCCs etc.)
- The whole content of the project website will be integrated into the European Language Grid website (proper) by the end of 2021



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

- The actual European Language Grid platform and website
- Functional ELG catalogue with more than a 1000 organisation entries

# Summary and Next Steps

- Establish ELG as the primary platform and marketplace for Language Technology in Europe.
- An initiative *from* the European LT community *for* the European LT community.
- European LT landscape is **highly fragmented**: ELG aims to provide just the right **umbrella platform**.
- Global market size by 2025 is enormous: we want the European LT community to be a **key player**.
- We want to **increase the visibility and reach** of all members of the European LT landscape.
- ELG is a long-term initiative: we will establish a **legal entity** for sustainability.
- Contribute to the long-term goal of **Digital Language Equality** in Europe by giving all our languages one virtual home and umbrella platform that collects **all** services and resources (**ELE**).
- **Next steps**: develop and populate **ELG Release 2** further with more functional services and resources; develop **ELG Release 3** (early 2022); establish the **ELG legal entity** (late 2021/early 2021).



European Language Equality

## The ELE project



The European Language Equality project has received funding from the European Union under grant agreement № LC-01641480 – 101018166 (ELE).

# European Language Equality: Consortium

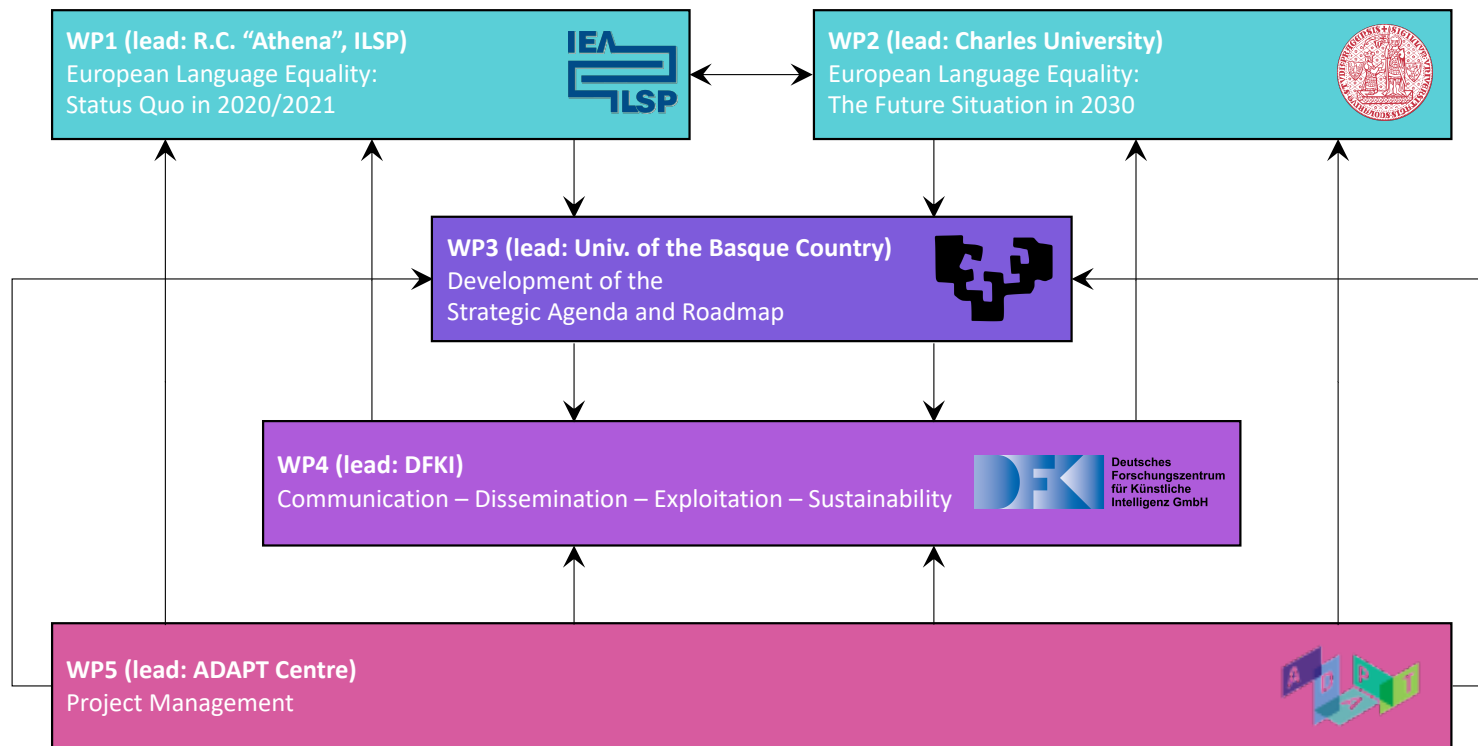
- **5 core partners:** Adapt Centre, DFKI, Charles University, ILSP, University of the Basque Country
- **9 networks, associations, initiatives:** LT Innovate (via Crosslang), EFNIL, ELEN, ECSPM, CLARIN, CLAIRE (via University of Leiden), NEM (via Eurescom), LIBER, Wikimedia
- **9 companies:** Tilde, ELDA, Expert System, Sail Labs, Kantan MT (via Xcelerator MT), Pangeanic, Semantic Web Company, Ontotext (via Sirma AI), SAP
- **29 research organisations:** University of Vienna, University of Antwerp, Institute for Bulgarian Language, University of Zagreb, University of Copenhagen, University of Tartu, University of Helsinki, CNRS, Research Institute for Linguistics, Institute for Icelandic Studies, FBK, University of Latvia, Institute of the Lithuanian Language, Luxembourg Institute of Technology, University of Malta, University of Utrecht, Language Council of Norway, Polish Academy of Sciences, University of Lisbon, Romanian Academy, University of Cyprus, Slovak Academy of Sciences, Jozef Stefan Institute, Barcelona Supercomputing Center, Royal Institute of Technology, University of Zurich, University of Sheffield, University of Vigo, Bangor University

# European Language Equality: Approach

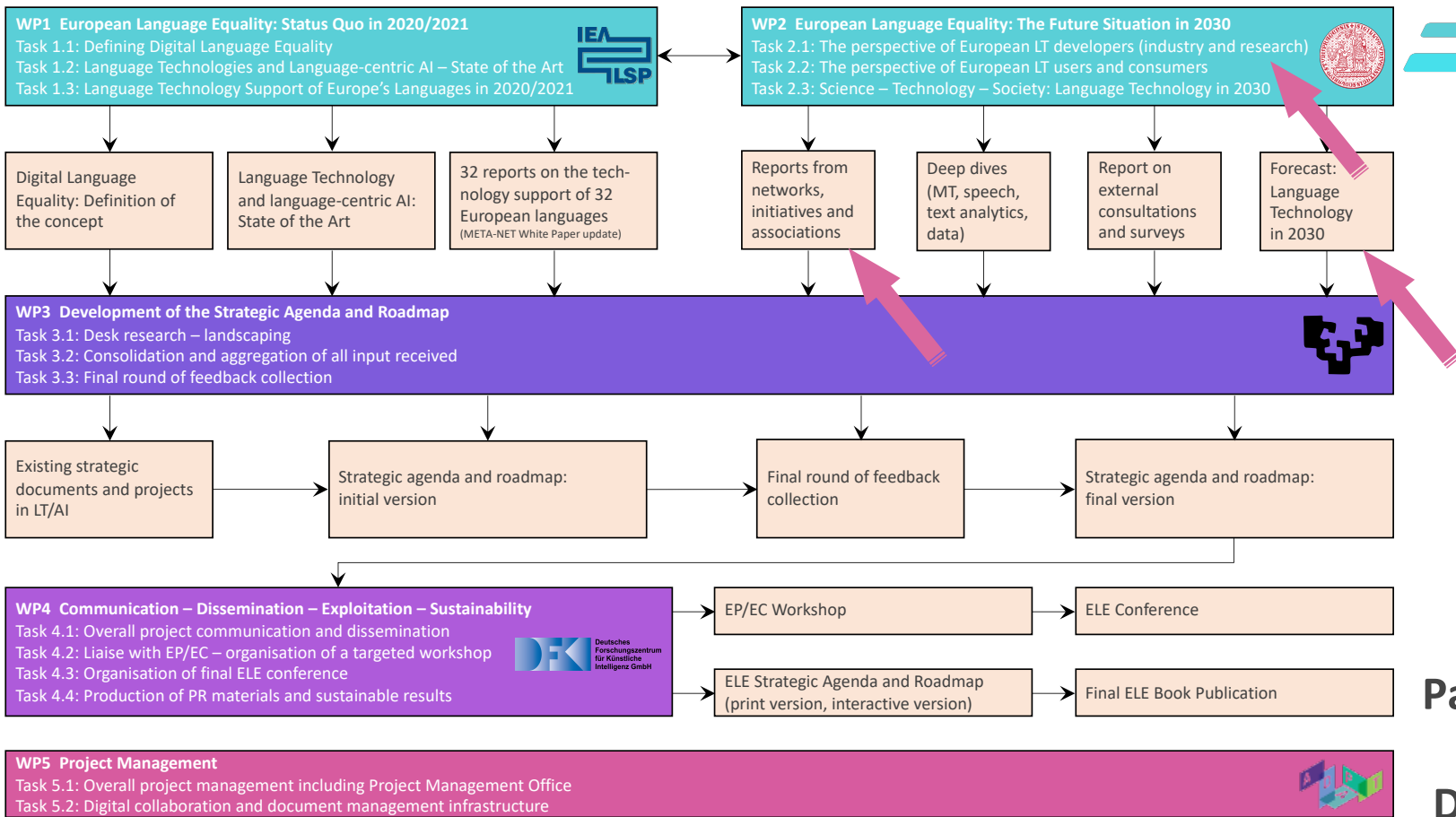
- Main result: Strategic Agenda and Roadmap – all deliverables provide input for this main report
  - Detailed description of the *European Language Equality Programme* (cf. EP resolution)
- *Research partners* prepare updates of META-NET White Papers (one deliverable each).
- *Networks and initiatives* to produce reports (one deliverable each) in which they collect, consolidate and present their own position, needs, wishes, demands, visions etc.
- *Companies* to produce various technical deep dives for the different technology areas.
- Several additional reports to be produced, primarily by the core partners.
- Reports to be prepared in a more or less autonomous way based on templates
- Reports to be used as input for the strategic agenda and roadmap – the main project result.
- Total number of languages taken into account in ELE: approx. 75.
- Close collaboration with ELG

META-NET

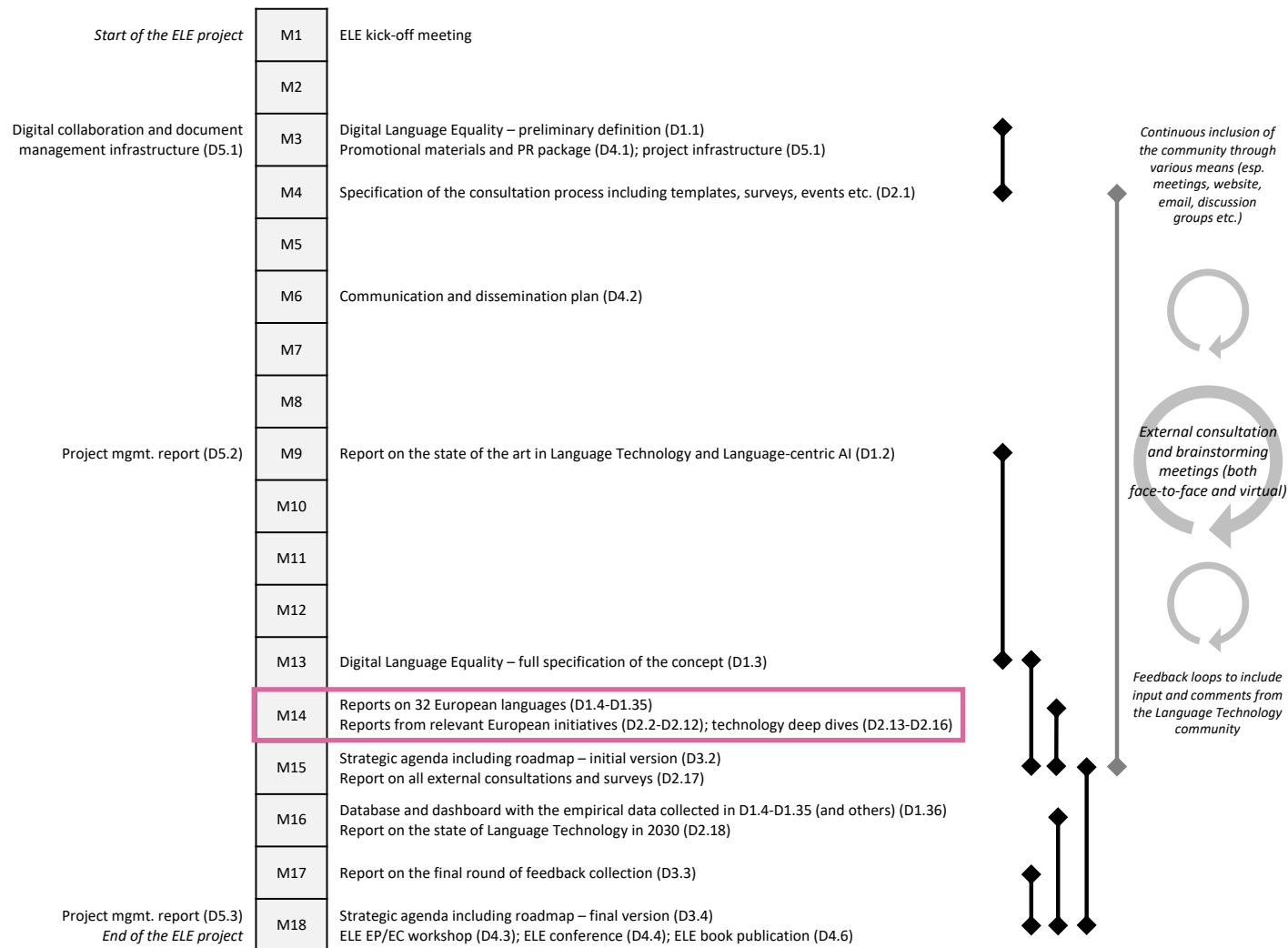




## Work Packages







# Timeline

<b>Network – Initiative – Association</b>	<b>Represented by ELE consortium partner(s)</b>
Association of European Research Libraries (LIBER)	LIBER
Big Data Value Association	SAP
Confederation of Laboratories for AI Research in Europe (CLAIRE)	ULEID
Cracking the Language Barrier	DFKI and various others
European Civil Society Platform for Multilingualism (ECSPM)	ECSPM
European Federation of National Institutions for Language (EFNIL)	EFNIL
European Language Equality Network (ELEN)	ELEN
European Language Grid (ELG)	DFKI, ILSP, CUNI, ELDA and others
European Lexicographic Infrastructure (ELEXIS)	JSI
European Research Infrastructure for LRs and Technology (CLARIN)	CLARIN ERIC (CUNI and ILSP are members)
LT-Innovate – Europe's LT Business Association	CRSLNG (SAIL, EXPSYS, TILDE are members)
META-NET	DFKI, CUNI, ILSP, TILDE, ELDA and others
New European Media (NEM)	ERSCM
Public-Private Partnership on AI (AI PPP)	SAP, TILDE
Wikipedia, Wikidata, Abstract Wikipedia	WMD
<b>External networks, initiatives and associations that ELE will consult with through established connections</b>	
AI4EU (European AI on Demand Platform), Dbpedia, Europeana, European Association for Machine Translation (EAMT), European Commission (DG Translate, DG Interpretation/SCIC), European Parliament (DG Translation, CULT Committee, ITRE Committee), Global WordNet Association (GWA), HumanE-AI-Net (and other ICT-48-2020 projects), Network to promote linguistic diversity (NPLD), World Wide Web Consortium (W3C) <i>and various others</i>	

## Networks and initiatives

No.	Deliverable name	WP	Short name	Type	Diss. level	Date
D1.1	Digital Language Equality – preliminary definition	1	DCU	R	Public	3
D1.2	Report on the state of the art in Language Technology and Language-centric AI	1	EHU	R	Public	9
D1.3	Digital Language Equality – full specification of the concept	1	DCU	R	Public	13
D1.4	Report on Basque	1	EHU	R+OTH	Public	14
D1.5	Report on Bulgarian	1	IBL	R+OTH	Public	14
D1.6	Report on Catalan	1	BSC	R+OTH	Public	14
D1.7	Report on Croatian	1	FFZG	R+OTH	Public	14
D1.8	Report on Czech	1	CUNI	R+OTH	Public	14
D1.9	Report on Danish	1	UCPH	R+OTH	Public	14
D1.10	Report on Dutch	1	UU	R+OTH	Public	14
D1.11	Report on English	1	USFD	R+OTH	Public	14
D1.12	Report on Estonian	1	UTART	R+OTH	Public	14
D1.13	Report on Finnish	1	UHEL	R+OTH	Public	14
D1.14	Report on French	1	CNRS	R+OTH	Public	14
D1.15	Report on Galician	1	UVIGO	R+OTH	Public	14
D1.16	Report on German	1	DFKI	R+OTH	Public	14
D1.17	Report on Greek	1	ILSP	R+OTH	Public	14
D1.18	Report on Hungarian	1	NYTI	R+OTH	Public	14
D1.19	Report on Icelandic	1	SAM	R+OTH	Public	14
D1.20	Report on Irish	1	DCU	R+OTH	Public	14
D1.21	Report on Italian	1	FBK	R+OTH	Public	14
D1.22	Report on Latvian	1	IMCS	R+OTH	Public	14
D1.23	Report on Lithuanian	1	LKI	R+OTH	Public	14
D1.24	Report on Luxembourgish	1	LIST	R+OTH	Public	14
D1.25	Report on Maltese	1	UOM	R+OTH	Public	14

**Deliverables**  
**1/3**

D1.26	Report on Norwegian	1	LCNOR	R+OTH	Public	14
D1.27	Report on Polish	1	PAS	R+OTH	Public	14
D1.28	Report on Portuguese	1	ULIS	R+OTH	Public	14
D1.29	Report on Romanian	1	ICIA	R+OTH	Public	14
D1.30	Report on Serbian	1	FILFUB	R+OTH	Public	14
D1.31	Report on Slovak	1	JULS	R+OTH	Public	14
D1.32	Report on Slovenian	1	JSI	R+OTH	Public	14
D1.33	Report on Spanish	1	BSC	R+OTH	Public	14
D1.34	Report on Swedish	1	KTH	R+OTH	Public	14
D1.35	Report on Welsh	1	BNGR	R+OTH	Public	14
D1.36	Database and dashboard with the empirical data collected in D1.4-D1.35 (and others)	1	ILSP	R+OTH	Public	16
D2.1	Specification of the consultation process including templates, surveys, events etc.	2	CUNI	R	Public	4
D2.2	Report from CLAIRE	2	ULEID	R+OTH	Public	14
D2.3	Report from CLARIN	2	CLARIN	R+OTH	Public	14
D2.4	Report from LT Innovate	2	CRSLNG	R+OTH	Public	14
D2.5	Report from META-NET	2	CUNI	R+OTH	Public	14
D2.6	Report from ELG	2	DFKI	R+OTH	Public	14
D2.7	Report from ECSPM	2	ECSPM	R+OTH	Public	14
D2.8	Report from EFNIL	2	EFNIL	R+OTH	Public	14
D2.9	Report from ELEN	2	ELEN	R+OTH	Public	14
D2.10	Report from LIBER	2	LIBER	R+OTH	Public	14
D2.11	Report from NEM	2	ERSCM	R+OTH	Public	14
D2.12	Report from Wikipedia	2	WMD	R+OTH	Public	14
D2.13	Technology deep dive Machine Translation	2	TILDE	R	Public	14
D2.14	Technology deep dive Speech Technologies	2	SAIL	R	Public	14
D2.15	Technology deep dive Text Analytics and Natural Language Understanding	2	EXPSYS	R	Public	14

**Deliverables  
2/3**

D2.16	Technology deep dive Data	2	SWC	R	Public	14
D2.17	Report on all external consultations and surveys	2	CUNI	R	Public	15
D2.18	Report on the state of Language Technology in 2030	2	CUNI	R	Public	16
D3.1	Report on existing strategic documents and projects in LT/AI	3	EHU	R	Public	3
D3.2	Strategic agenda including roadmap – initial version	3	DFKI	R	Public	15
D3.3	Report on the final round of feedback collection	3	EHU	R	Public	17
D3.4	Strategic agenda including roadmap – final version	3	DFKI	R	Public	18
D4.1	Promotional materials and PR package	4	DFKI	R+OTH	Public	3
D4.2	Communication and dissemination plan	4	DFKI	R	Public	6
D4.3	Report on EP/EC workshop	4	DFKI	R	Public	18
D4.4	Report on ELE conference	4	DFKI	R	Public	18
D4.5	Strategic agenda and roadmap (print version, online version)	4	DFKI	R	Public	18
D4.6	ELE book publication	4	DFKI	R	Public	18
D5.1	Digital collaboration and document management infrastructure	5	DCU	R+OTH	Confidential	3
D5.2	Project management report (interim report)	5	DCU	R	Confidential	9
D5.3	Project management report (final report)	5	DCU	R	Confidential	18

# Deliverables 3/3

# Summary

- ELE is a new EU project that started in January 2021 and ends in June 2022
- Its goal is the development of the *Strategic Research, Innovation and Implementation Agenda and a Roadmap for achieving full Digital Language Equality in Europe by 2030*
- Once-in-a-decade opportunity
- Consortium with 52 partners covering all European countries and all major initiatives
- Many consultation events, roundtables, stakeholder meetings planned
- Close collaboration with its sister project European Language Grid (ELG)
- Intended overlap between ELE and ELG consortium, ELG NCCs and META-NET members
- Firmly establish LT and language-centric AI in Horizon Europe and Digital Europe Programme
- Please participate in our stakeholder events, surveys and questionnaires!
- ELE and ELG will finish up with a joint META-FORUM 2022 in June 2022.

# Developing an agenda and a roadmap for achieving full digital language equality in Europe by 2030



## Overview

With a large and all-encompassing consortium consisting of 52 partners covering all European countries, research and industry and all major pan-European initiatives, the European Language Equality (ELE) project develops a strategic research, innovation and implementation agenda as well as a roadmap for achieving full digital language equality in Europe by 2030. The whole project is designed in such a way that all individual work streams create input for the strategic agenda, for example, in terms of empirical evidence, desk re-

## Results

The main tangible results to be produced by the ELE project are as follows:

- The research partners will prepare updates of the 32 META-NET White Papers, which will also include the results of an enormous and exhaustive data collection procedure that will provide a detailed, empirical and dynamic map of technology support for our languages.
- A number of networks and initiatives (CLAIRE, CLARIN, LT Innovate, META-NET, ELG, ECSPM, EFNIL ELEN, LIBER, NEM and

## Motivation

The ELE project is a direct response to the resolution “Language equality in the digital age”, which was passed by the European Parliament in a landslide vote – **592 votes in favour and only 45 against – in September 2018. The report provides more than 40 recommendations, including:**

25. Establish a large-scale, long-term coordinated funding programme for research, development and innovation in the field of language technologies, at European, national and regional levels, tailored specifically to Eu-

You're invited to stay in touch via <https://european-language-equality.eu>

# Thank you!



The European Language Grid has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement № 825627 (ELG).



The European Language Equality project has received funding from the European Union under grant agreement № LC-01641480 – 101018166 (ELE).

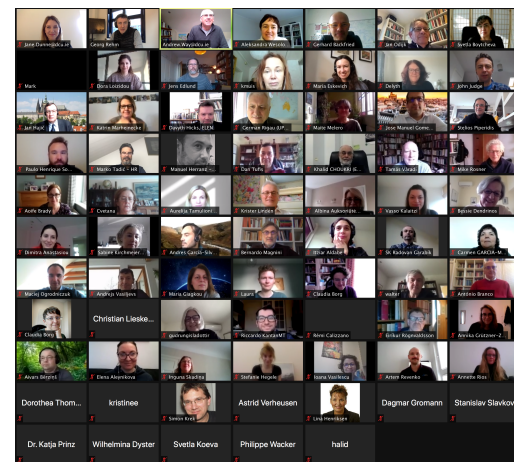
Coordinator ELG and Co- Coordinator ELE: Georg Rehm (DFKI GmbH)

Coordinator ELE: Andy Way, ADAPT Centre (DCU)

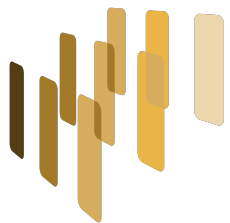
02-06-2021 Latvian Conference “Digital resources and tools for education”

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

<https://www.european-language-grid.eu>







# EUROPEAN LANGUAGE GRID

**Interested  
to learn  
more?**

- Visit our project website:

<https://www.european-language-grid.eu>

- Go to the live ELG:

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

- Like to interact with ELG? Send an email to:

[contact@european-language-grid.eu](mailto:contact@european-language-grid.eu)