EUROPEAN LANGUAGE EQUALITY

EUROPEAN LANGUAGE GRID

European Language Equality: An Overview

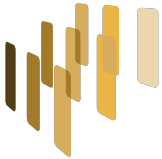
http://www.european-language-equality.eu

Presenter: Guðrún Gísladóttir ECSPM Project LEAR (Legal Entity Appointed Representative)

Using a slightly adapted PPT created by Georg Rehm (Co-Coordinator ELE, DFKI)

European Language Equality (ELE) – Summary

EUROPEAN LANGUAGE EQUALITY



EUROPEAN LANGUAGE GRID

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

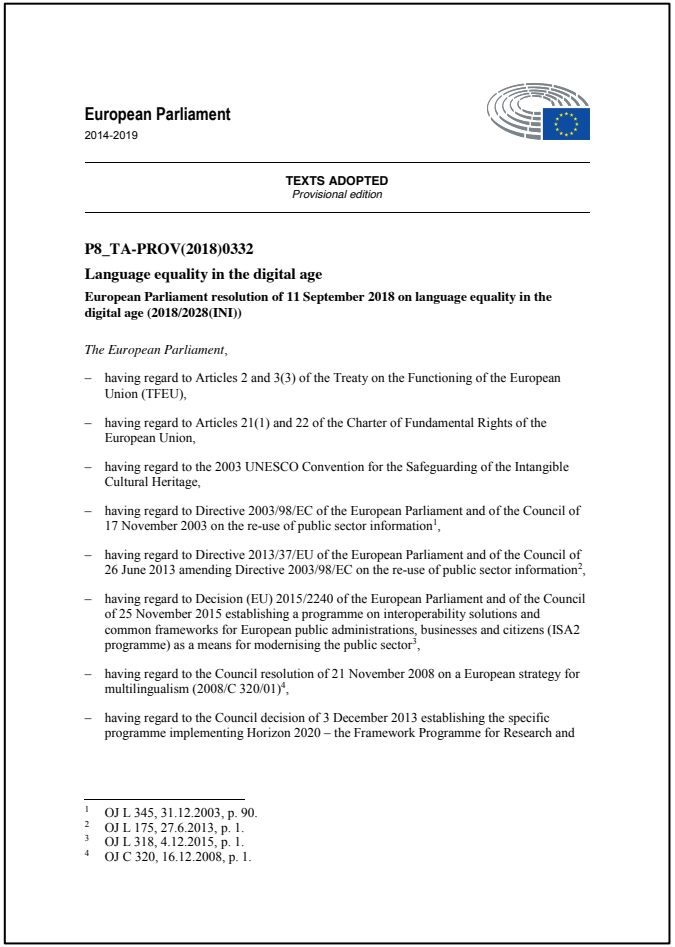
Start time: 1 January 2021

Completion: June 2022

Context: “Language Equality” EP Resolution

This flagship project addresses recommendations in the 2018 European Parliament Resolution **Language equality in the digital age**

- 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
- Create a European LT platform for sharing of services
- Europe has to secure its leadership in language-centric AI; recalls that EU companies are the best placed to provide solutions tailored to our specific cultural, societal and economic needs.



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[1],

– 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[2],

– 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[3],

– having regard to the Council resolution of 21 November 2008 on a European strategy for multilingualism (2008/C 320/01)[4],

– having regard to the Council decision of 3 December 2013 establishing the specific programme implementing Horizon 2020 – the Framework Programme for Research and

[1] OJ L 345, 31.12.2003, p. 90.
[2] OJ L 175, 27.6.2013, p. 1.
[3] OJ L 318, 4.12.2015, p. 1.
[4] OJ C 320, 16.12.2008, p. 1.

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, Jožef Stefan Institute, Barcelona Supercomputing Center, Royal Institute of Technology, University of Zurich, University of Sheffield, University of Vigo, Bangor University

European Language Equality: Approach

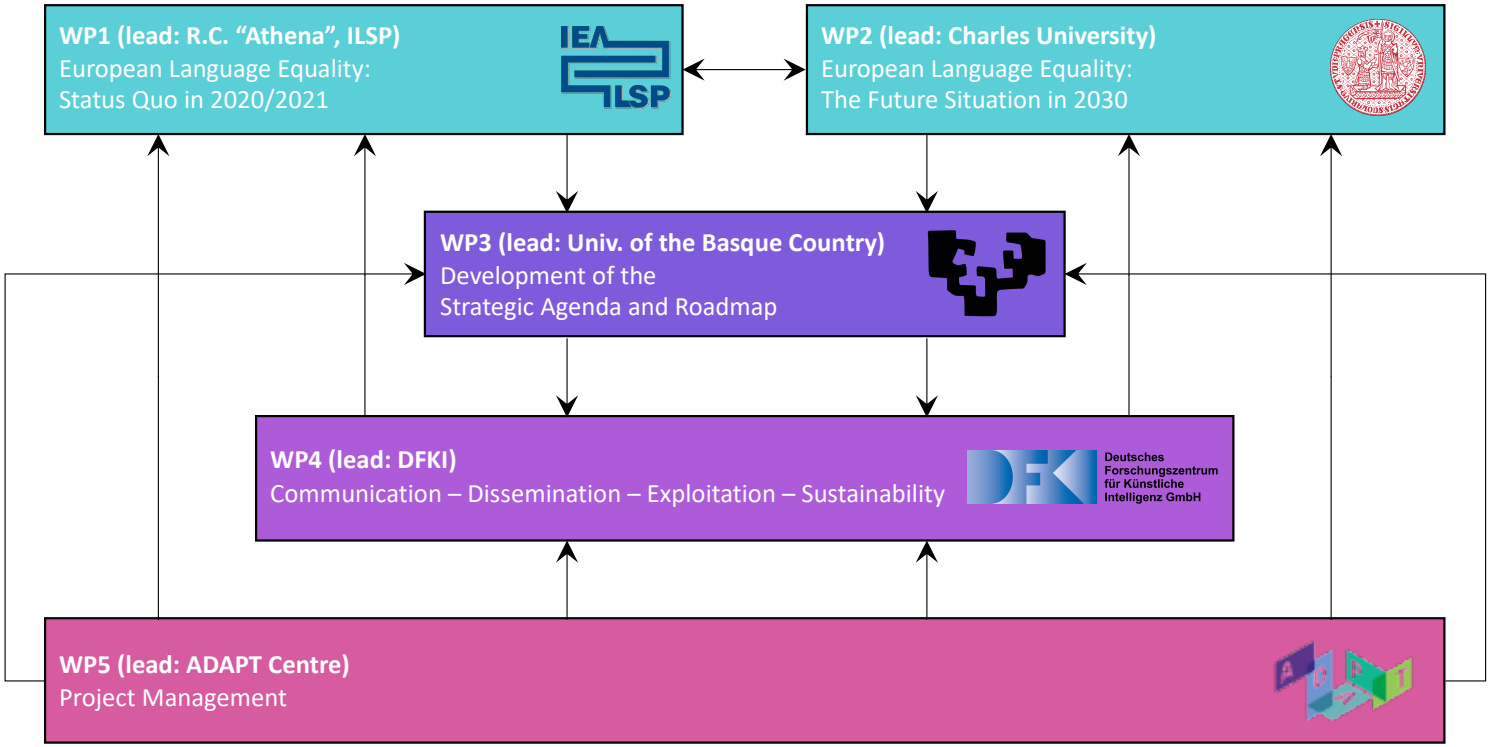
- Main result: **Strategic Agenda and Roadmap** (all deliverables provide input for this main report)

Networks and initiatives

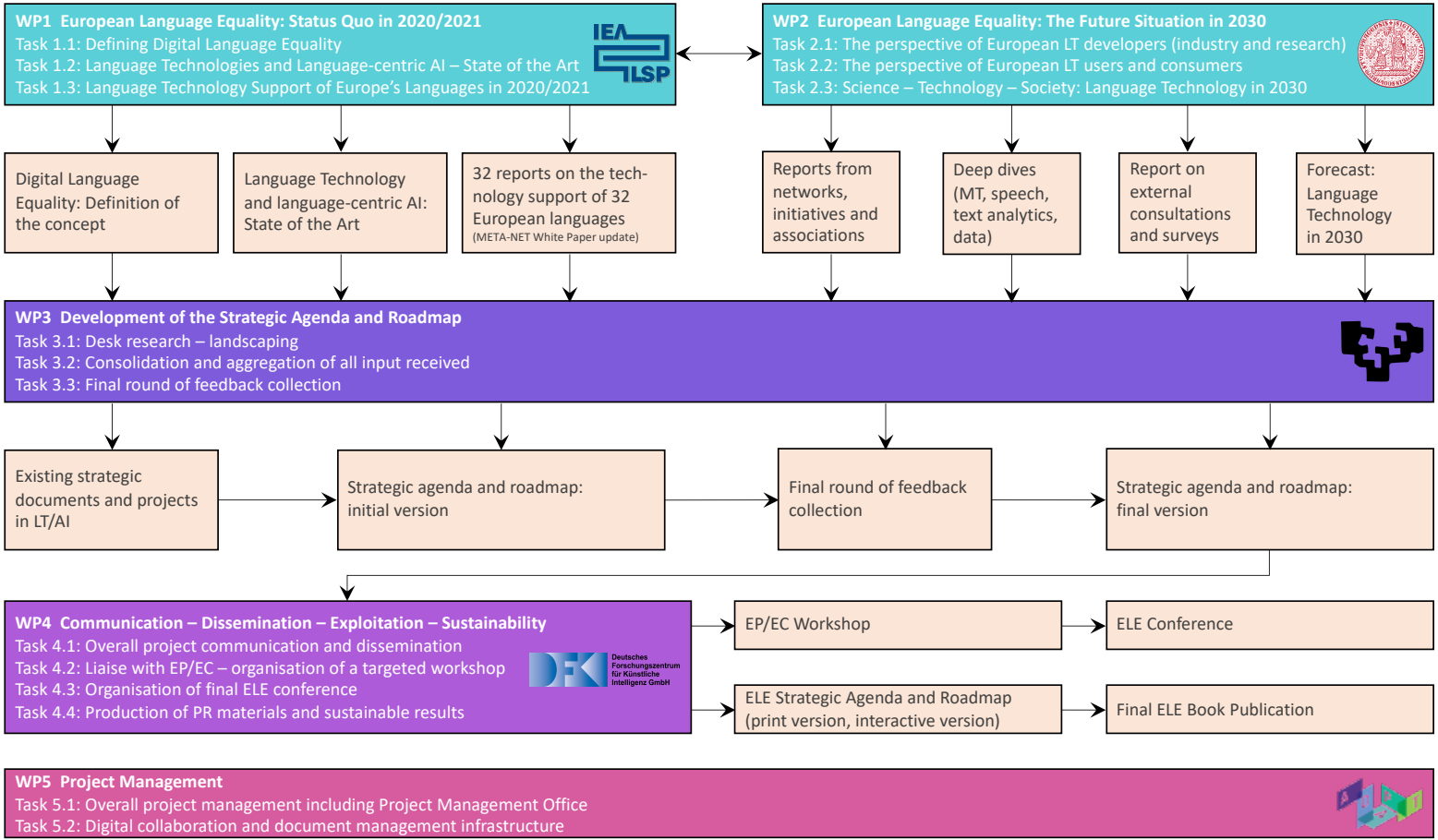
- To produce reports (one deliverable each) in which they collect, consolidate and present their own position, needs, wishes, demands, visions etc.
- Companies to produce various technical deep dives for the different technology areas
- Several additional reports to be produced, primarily by the core 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**.

Throughout the project there is

- Close collaboration with ELG



Work Packages

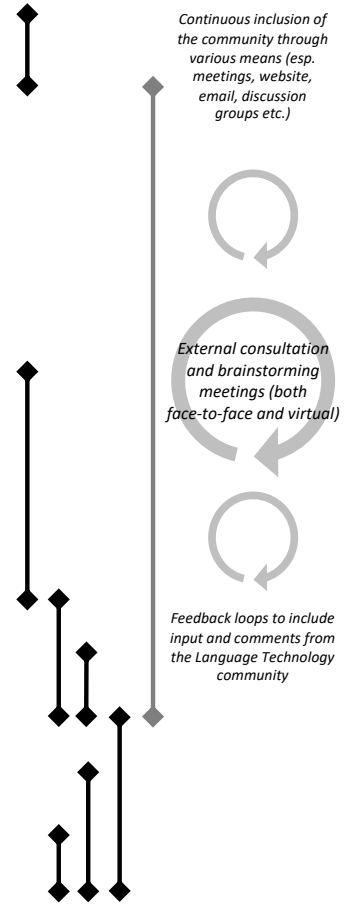


EUROPEAN LANGUAGE EQUALITY

Work Packages and main Deliverables

Timeline

Start of the ELE project	M1	ELE kick-off meeting
	M2	
Digital collaboration and document management infrastructure (D5.1)	M3	Digital Language Equality – preliminary definition (D1.1) Promotional materials and PR package (D4.1); project infrastructure (D5.1)
	M4	Specification of the consultation process including templates, surveys, events etc. (D2.1)
	M5	
	M6	Communication and dissemination plan (D4.2)
	M7	
	M8	
Project mgmt. report (D5.2)	M9	Report on the state of the art in Language Technology and Language-centric AI (D1.2)
	M10	
	M11	
	M12	
	M13	Digital Language Equality – full specification of the concept (D1.3)
	M14	Reports on 32 European languages (D1.4-D1.35) Reports from relevant European initiatives (D2.2-D2.12); technology deep dives (D2.13-D2.16)
	M15	Strategic agenda including roadmap – initial version (D3.2) Report on all external consultations and surveys (D2.17)
	M16	Database and dashboard with the empirical data collected in D1.4-D1.35 (and others) (D1.36) Report on the state of Language Technology in 2030 (D2.18)
	M17	Report on the final round of feedback collection (D3.3)
Project mgmt. report (D5.3) *End of the ELE project*	M18	Strategic agenda including roadmap – final version (D3.4) ELE EP/EC workshop (D4.3); ELE conference (D4.4); ELE book publication (D4.6)



Network – Initiative – Association	Represented by ELE consortium partner(s)
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
External networks, initiatives and associations that ELE will consult with through established connections	
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) *and various others*	

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

EUROPEAN LANGUAGE EQUALITY

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

EUROPEAN LANGUAGE EQUALITY

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

EUROPEAN LANGUAGE EQUALITY

Contribution to the ELE by the ECSPM

ECSPM Member organisations and languages

Cooperation with three of our member organisations:

- **FUEN** (Federal Union of European Nationalities) are primary informants on:
  Aromanian (250.000), **Carpathian-German** (4.690), **Carpato-Rusyn** (69.060), **Lezghin** (800.000), **Meskhetian** (Ahiska) (550.000), **Moldovian** (2.030.000), **Pomak** (600.000) and **Roma** (3.500.000)
  Secondary informants: Czech and Kashubian
- **NPLD** (Network to Promote Linguistic Diversity) are primary informants on:
  Ladin (30.000) and **Macedonian** (2.200.000)
  Secondary informants: Breton, Catalan, Frisian, Friulian, Occitan, Sardinian
- **CEM** (Centre of Excellence for Multilingualism and Language Policy, NKUA)
  Turkish (76.700.000)
- **ECSPM** directly
  Albanian (7.500.000)

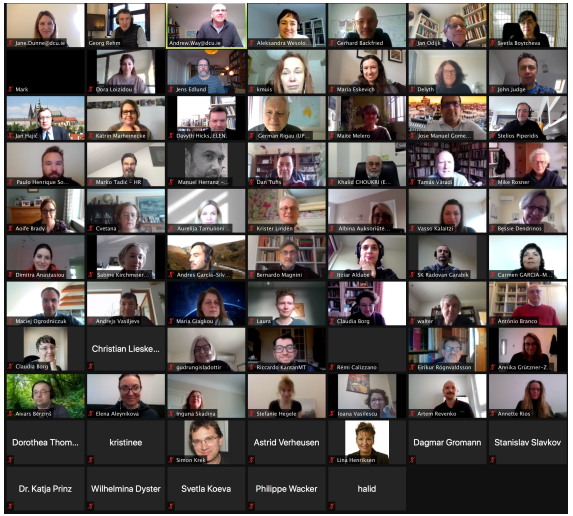
ECSPM's contribution

- So far:
  - Contributed to defining terminology (Official, Regional, Autochthonous/indigenous minority languages, Community languages, Heritage languages)
  - Contributed to the definition of Digital Language Equality
  - Discovered, collected and documented data sets, tools, repositories etc. for the languages mentioned on the previous slide
  - Helped create a digital future for the smaller languages. Currently, for quite a few, there are very limited resources, they are simply in the danger of digital extinction
  - Have put Turkish and Albanian, which are big European heritage languages, on the digital map
- Next steps:
  - To keep ECSPM and ECSPM-CURUM members informed
  - To exchange ideas on how to benefit from the project

European Language Equality

Bestu þakkir – Thank you

Tack – Ευχαριστώ πολύ



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

The ELE website:

https://european-language-equality.eu