

EUROPEAN LANGUAGE GRID



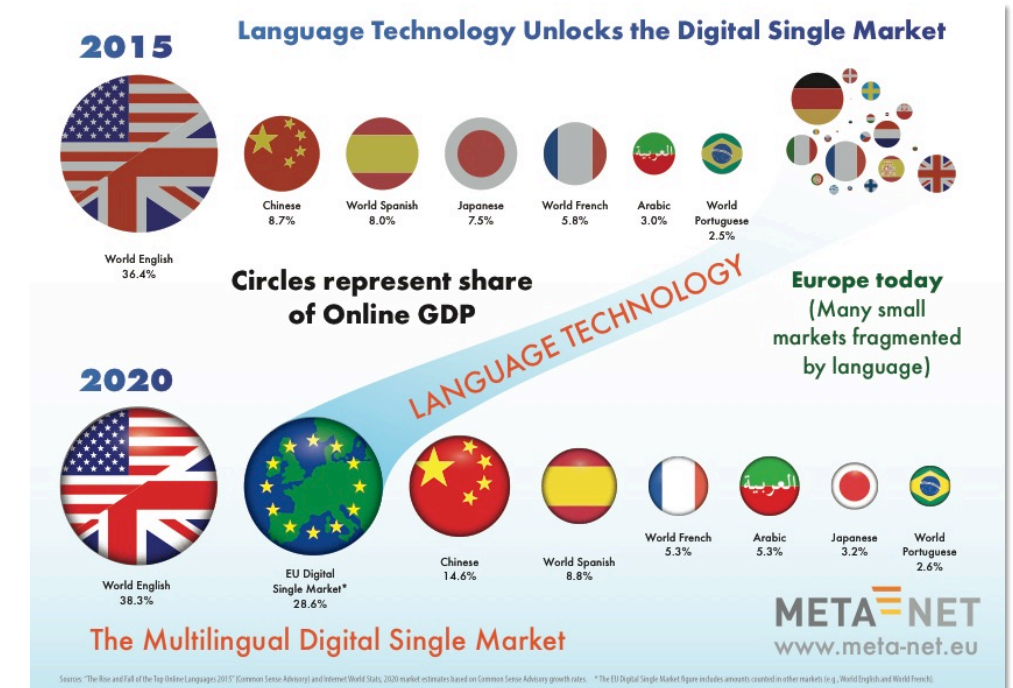
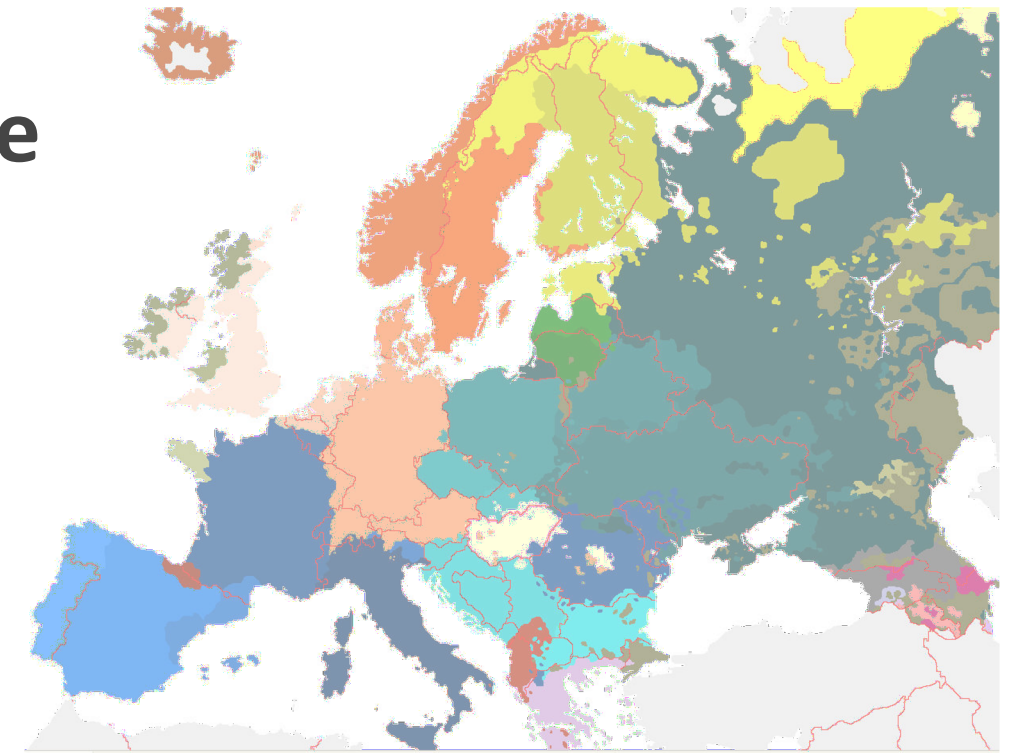
The European Language Grid and European Language Equality Projects

Prof. Dr. Georg Rehm (DFKI GmbH, Germany)
Coordinator ELG, Co-coordinator ELE

16-11-2021 ELG NCC Workshop Denmark – Danish Language Technology Conference 2021
<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
 - The Digital Single Market needs to be multilingual
 - Cross-border, cross-lingual, cross-cultural communication
 - Fragmentation of the LT market and landscape





META FORUM 2010

T4ME

T4ME
EU-funded project (Seventh Framework Programme) working on technologies for the Multilingual European Information Society

META-NET
Established in 2010, META-NET is a network of Excellence consisting of 60 research centres from 34 countries, building the technological foundations of a multilingual European information society

META-FORUM 2010
"Challenges for Multilingual Europe" (November 17/18, 2010)



META SHARE

META FORUM 2012

META-FORUM 2012
Conference "A Strategy for Multilingual Europe" (Brussels – June 20/21, 2012)

META-NET White Papers
Release of 32 volumes on 31 languages, revealing that there is a severe threat of digital extinction for at least 21 European languages (December 2012)

META FORUM 2015



CRACKER
EU-funded project CRACKER (Horizon2020) pushing towards an improvement of MT research in terms of efficiency and effectiveness (2015 – 2017)

Cracking the Language Barrier Federation
Founded in 2015, the federation has been assembling European research and innovation projects as well as all related community organisations working on multilingual technologies

META-FORUM 2015
Conference "Technologies for the Multilingual Digital Single Market" (Riga – April 27, 2015)

Strategic Agenda for the Multilingual Digital Single Market (Version 0.5)
Launch of the Strategic Agenda for the Multilingual Digital Single Market titled "Technologies for Overcoming Language Barriers towards a truly integrated European Online Market" (April 2015)

Riga Summit on the Multilingual Digital Single Market
Summit "Shape the future of the multilingual digital single market" (April 27–29, 2015)

META FORUM 2017

"Language Equality in the Digital Age"
Workshop on "Language Equality in the Digital Age", commissioned by the EU Parliament's Science and Technology Options Assessment Committee (STOA) (January 2017)

"Language equality in the digital age: Towards a Human Language Project"
Launch of the study on "Language equality in the digital age: Towards a Human Language Project", commissioned by the EU Parliament (March 2017)

META-FORUM 2017
Conference "Towards a Human Language Project" (Brussels – November 13/14, 2017)

Strategic Research and Innovation Agenda (V1.0)
Launch of the Strategic Research and Innovation Agenda titled "Language Technologies for Multilingual Europe – Towards a Human Language Project" (December 2017)



2010 2011 2012 2013 2014 2015 2016 2017 2018

META NORD

META-NORD
EU-funded (ICT PSP) Project to establish an open linguistic infrastructure in the Baltic and Nordic countries (2011 – 2013)

METANET4U

METANET4U
EU-funded project (ICT Policy Support Programme) to enhance the European Linguistic Infrastructure (2011 – 2013)

CESAR

CESAR
EU-funded project (ICT Policy Support Programme) functioning as a part of META-NET to standardise language resources and tools (2011 – 2013)

META-FORUM 2011
Conference "Solutions for Multilingual Europe" (Budapest – June 27/28, 2011)

"State of the Art of Machine Translation – Current Challenges and Future Opportunities"
Workshop on "State of the Art of Machine Translation", commissioned by the EU Parliament (December 2013)

META-SHARE
Initiated in 2013, META-SHARE has functioned as an open and secure network of repositories for sharing and exchanging language data, tools and services

Strategic Research Agenda for Multilingual Europe 2020
Launch of the Strategic Research Agenda for Multilingual Europe 2020 (January 2013)

META-FORUM 2013
Conference "Connecting Europe for New Horizons" (Berlin – September 19/20, 2013)

Strategic Research and Innovation Agenda (Version 0.9)
Launch of the Strategic Research and Innovation Agenda titled "Language as a Data Type and Key Challenge for Big Data" (July 2016)

META-FORUM 2016
Conference "Beyond Multilingual Europe" (Lisbon – July 4/5, 2016)



EUROPEAN LANGUAGE GRID

2019-2022

EUROPEAN LANGUAGE EQUALITY

2021-2022

META FORUM 2011

META FORUM 2013

META FORUM 2016

ELG consortium



ELG European Language Grid



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:

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

27. Europe has to secure its leadership in language-centric AI

32. Set up LT financing platform; emphasise R&D in Deep NLU

40. Develop investment instruments and accelerator programs

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,

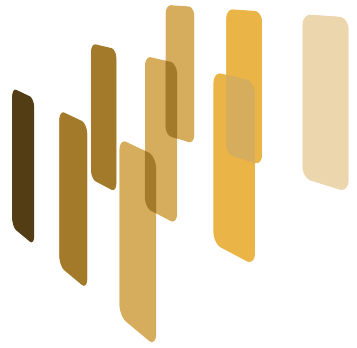
- having regard to Articles 2 and 3(3) of the Treaty on the Functioning of the European Union (TFEU),
- having regard to Articles 21(1) and 22 of the Charter of Fundamental Rights of the European Union,
- having regard to the 2003 UNESCO Convention for the Safeguarding of the Intangible Cultural Heritage,
- 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 information¹,
- 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 information²,
- 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 sector³,
- 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 GRID

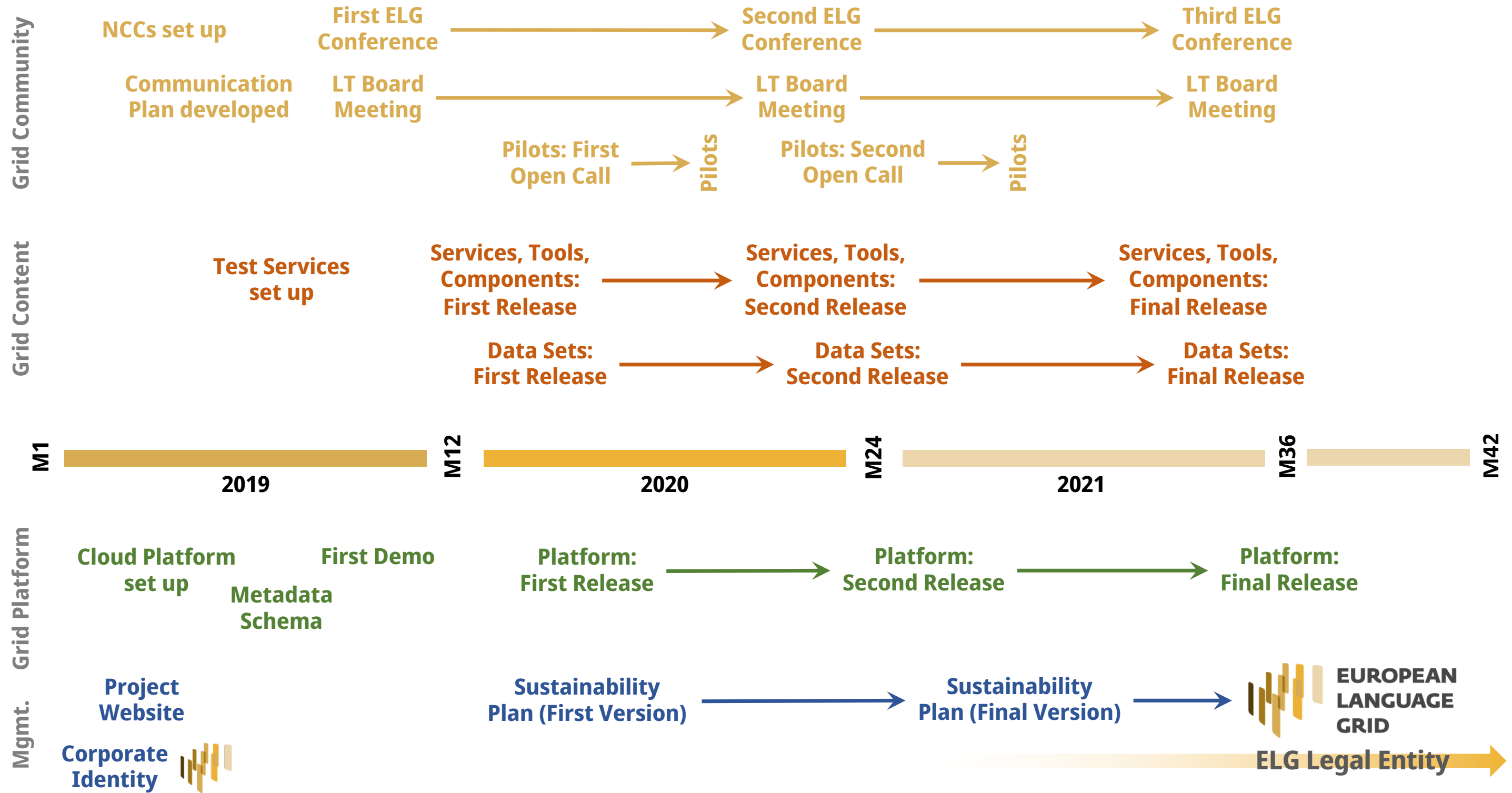
Objectives (Selection)

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.



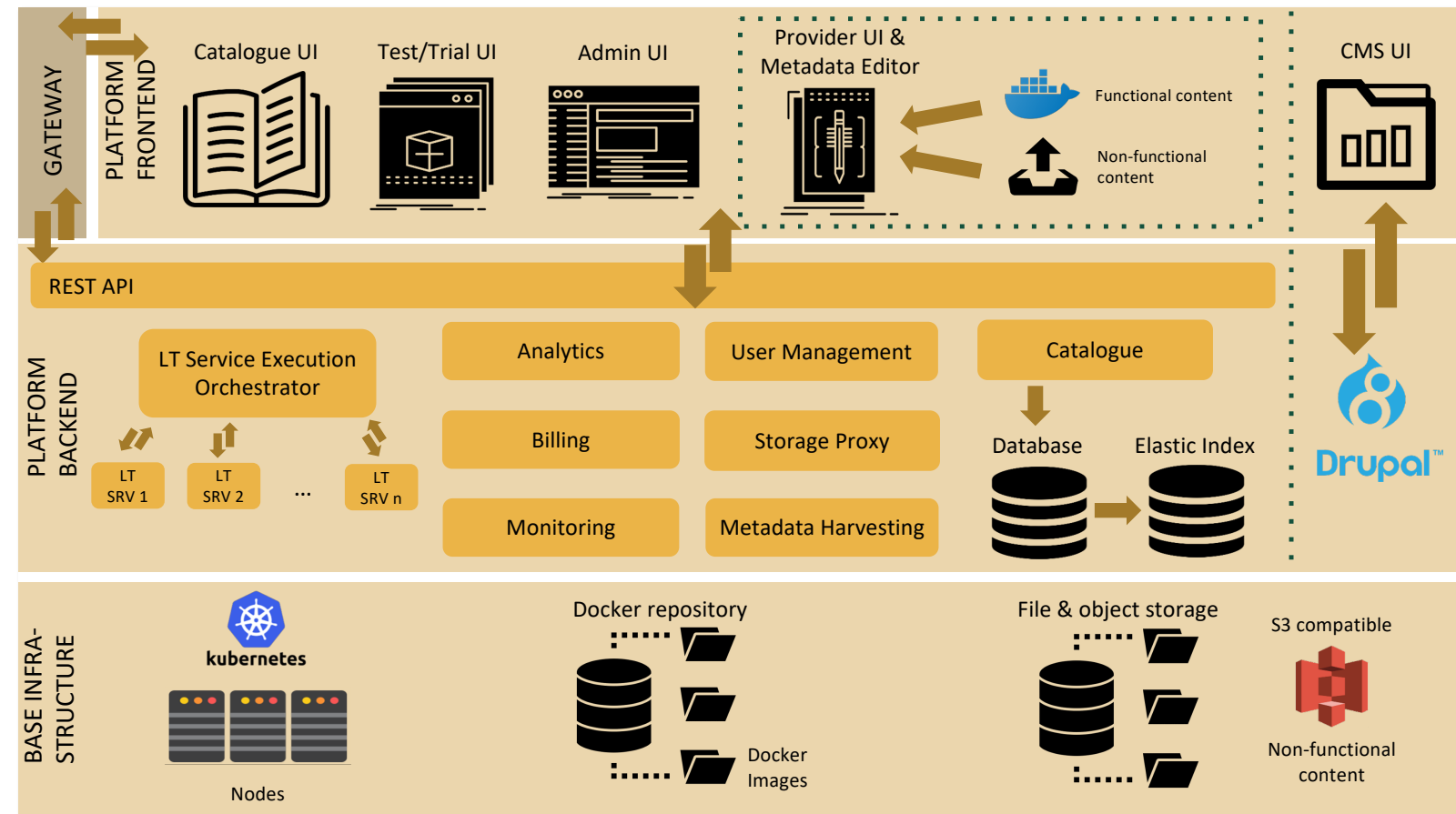
Kick-off meeting, 22/23 January 2019





European Language Grid – Current State of Play (November 2021)

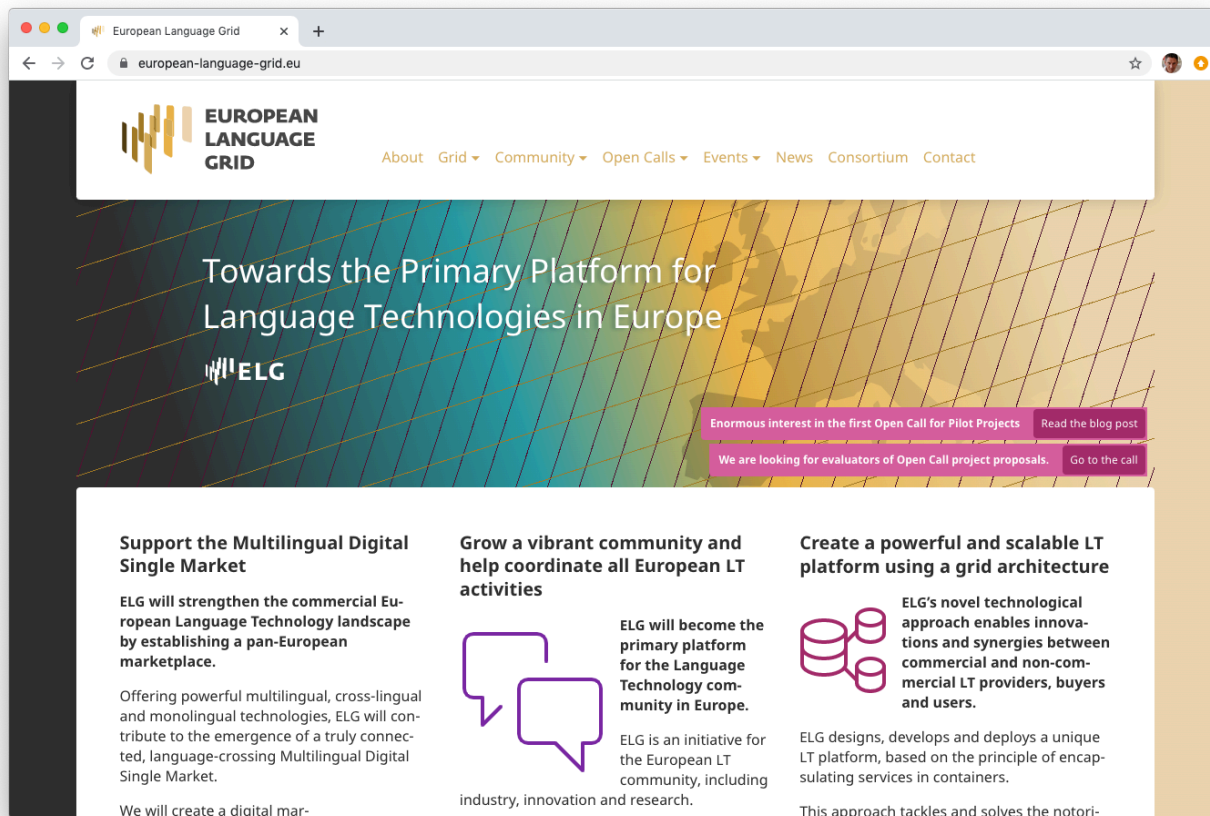
- User registration, authentication, authorisation
- User categories, respective rights and policies
- LT metadata upload
- Metadata conversion and harvesting
- LT service registration, integration
- LT service try out and execution
- LT data browsing, searching, downloading
- **Current state of play ELG Release 2 (Nov. 2021):**
 - 450 functional services and tools
 - 3563 corpora and data sets
 - 948 lexical/conceptual resources
 - 27 language descriptions, grammars, models
 - 1789 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.

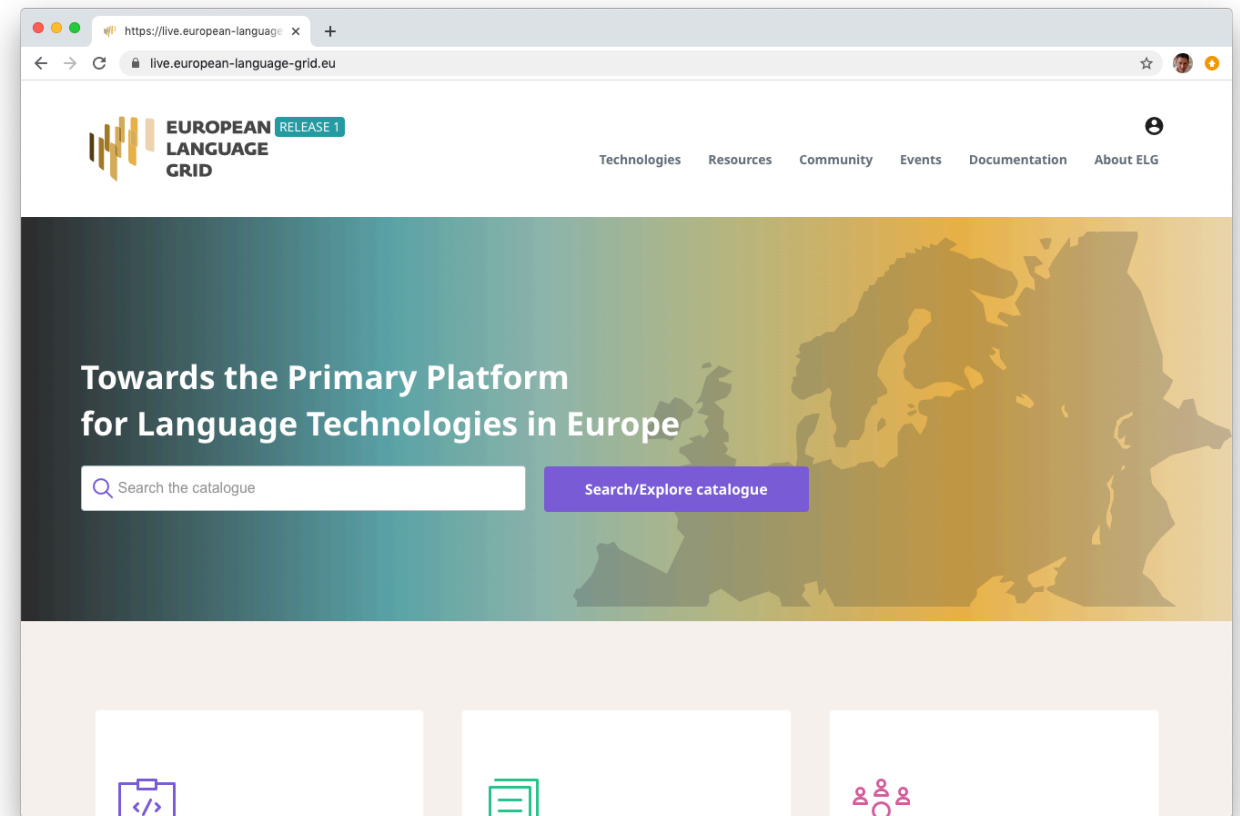
ELG Release 3 to be made available in early 2022.

European Language Grid – Quick Demo



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

- ELG project website
- Typical information about our EU project 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) in the second half of 2021



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

- The actual European Language Grid platform and website
- The next revision of the ELG project website (left) will include dynamic content that will be pulled from the ELG repository (right), e.g., languages, number of services, categories of services etc.



Georg Rehm

- Technologies
- Resources
- Community
- Events
- Documentation
- About ELG

Towards the Primary Platform for Language Technologies in Europe



Language technologies

LT services, tools, components, downloadable or deployed directly through the grid

[Browse Technologies](#)



Language data and resources

The collection of data sets, corpora, language models and other language resources

[View Resources](#)



Community

Organizations, projects, events etc in European Language technology field

[Explore Community](#)



Georg Rehm

- Technologies
- Resources
- Community
- Events
- Documentation
- About ELG

Towards the Primary Platform for Language Technologies in Europe



Language technologies

LT services, tools, components, downloadable or deployed directly through the grid

[Browse Technologies](#)



Language data and resources

The collection of data sets, corpora, language models and other language resources

[View Resources](#)



Community

Organizations, projects, events etc in European Language technology field

[Explore Community](#)



Georg Rehm

- Technologies
- Resources
- Community
- Events
- Documentation
- About ELG

UEDIN-MT-DeEn

Search



- Language resources & technologies
- Service functions
- Languages
- Licences
- Conditions of use for data
- Related entities

140 search results for UEDIN-MT-DeEn



UEDIN Machine Translation Service for German to English

1.0.0

111 views

2131 times used

ELG-compatible service

A machine translation (MT) service for German-to-English translation based on the Marian machine translation framework. The translation model is a basic transformer model trained on ca 13.3M sentence pairs using Marian N

Keywords: Machine Translation · German · English · Neural machine translation

Languages: German · English

Licence: Creative Commons Attribution Share Alike 4.0 International



UEDIN Machine Translation Service for Czech to English

1.0.0

25 views

152 times used

ELG-compatible service

An ensemble of 3 transformer-base models trained on WMT'19 data as a system for generating back-translations.

Keywords: Machine Translation · Czech · English · Neural machine translation

Languages: English · Czech

Licence: Creative Commons Attribution Share Alike 4.0 International



UEDIN Machine Translation Service for English to German

1.0.0

62 views

570 times used

ELG-compatible service

A machine translation (MT) service for English-to-German translation based on the Marian machine translation framework. The translation model is a basic transformer model trained on ca 13.3M sentence pairs using Marian N

Keywords: Machine Translation · German · English · Neural machine translation

Languages: German · English

Licence: Creative Commons Attribution Share Alike 4.0 International



Georg Rehm

- Technologies
- Resources
- Community
- Events
- Documentation
- About ELG



- conll shared task
- 2007 conll shared task
- hosted at elg
- udpipe english
- Service functions
- Languages
- Licences
- Conditions of use for data
- Related entities

"Fatalne jaja" Bułhakow

33 views
0 downloads

Story "Fatalne Jaja" Michaił Bułhakow

Keywords: korpus · korpus tekstowy

Language: Polish

Licence: Creative Commons Attribution 3.0 Unported

"Le Monde Diplomatique" Arabic tagged corpus

28612 views
0 downloads

This corpus contains 102,960 vowelised, lemmatised and tagged words (58 texts from Le Monde Diplomatique Arabic, see also ELRA-W0036-04). To each text are associated 3 files : - raw text in Arabic, - vowelized text in

Keyword: corpus

Language: Arabic

Licences: ELRA-VAR-COMMERCIAL-MEMBER-COMMERCIALUSE-1.0

· ELRA-END-USER-COMMERCIAL-MEMBER-NONCOMMERCIALUSE-1.0

"Le Monde Diplomatique" Text corpus in Arabic

116 views
0 downloads

Electronic archiving of "Le Monde Diplomatique" articles in Arabic from 2000. The corpus is available in HTML. Each HTML file contains one article. Number of articles available per year : • 2000: 61 articles (November a

Keyword: corpus

Language: Arabic

Licences: ELRA-END-USER-ACADEMIC-MEMBER-NONCOMMERCIALUSE-1.0

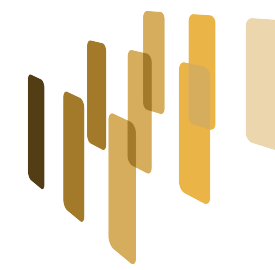
· ELRA-END-USER-COMMERCIAL-MEMBER-NONCOMMERCIALUSE-1.0

Some statistics on ingested LRs that are publicly visible

Repository	Corpora	Lexical/Conceptual Resources	Models & Computational grammars	Total
ELRA	635	545	–	1180
ELRC-SHARE	1249	50	–	1299
META-SHARE	52	12	7	71
ELRA-SHARE-LRs	92	33	1	126
LINDAT/CLARIAH-CZ	274	79	–	353
CLARIN.SI	140	78	–	218
CLARIN.PL	239	15	–	254
Quantum Stat	255	5	–	260
Zenodo	137	99	24	*260
Tudatalib	1	–	–	1
HuggingFace	385	–	–	385
Others	72	28	5	105
TOTAL	3531	944	37	4512

*approx. 400 ingested resources are still under validation

Resource Types

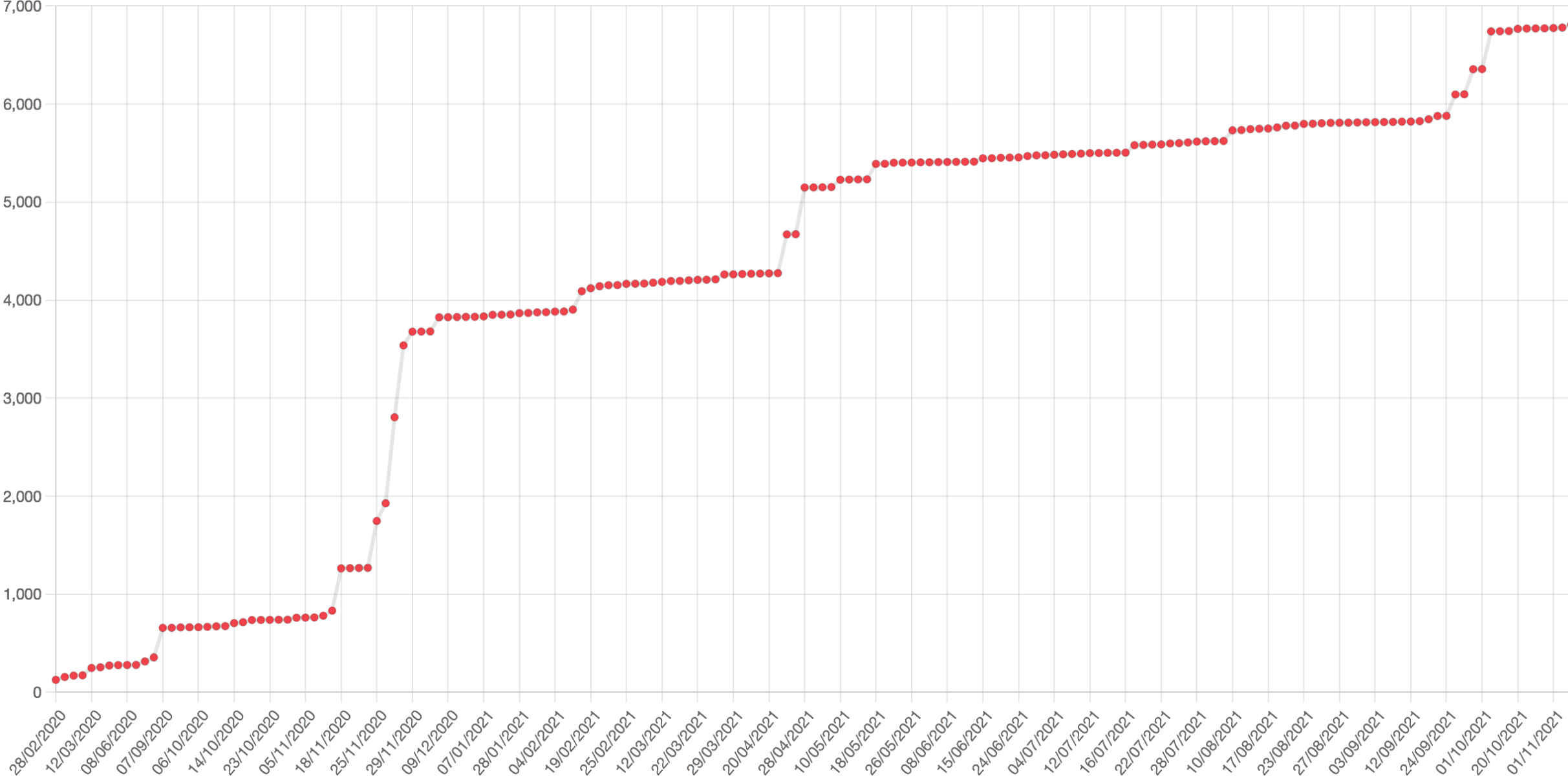


**EUROPEAN
LANGUAGE
GRID**

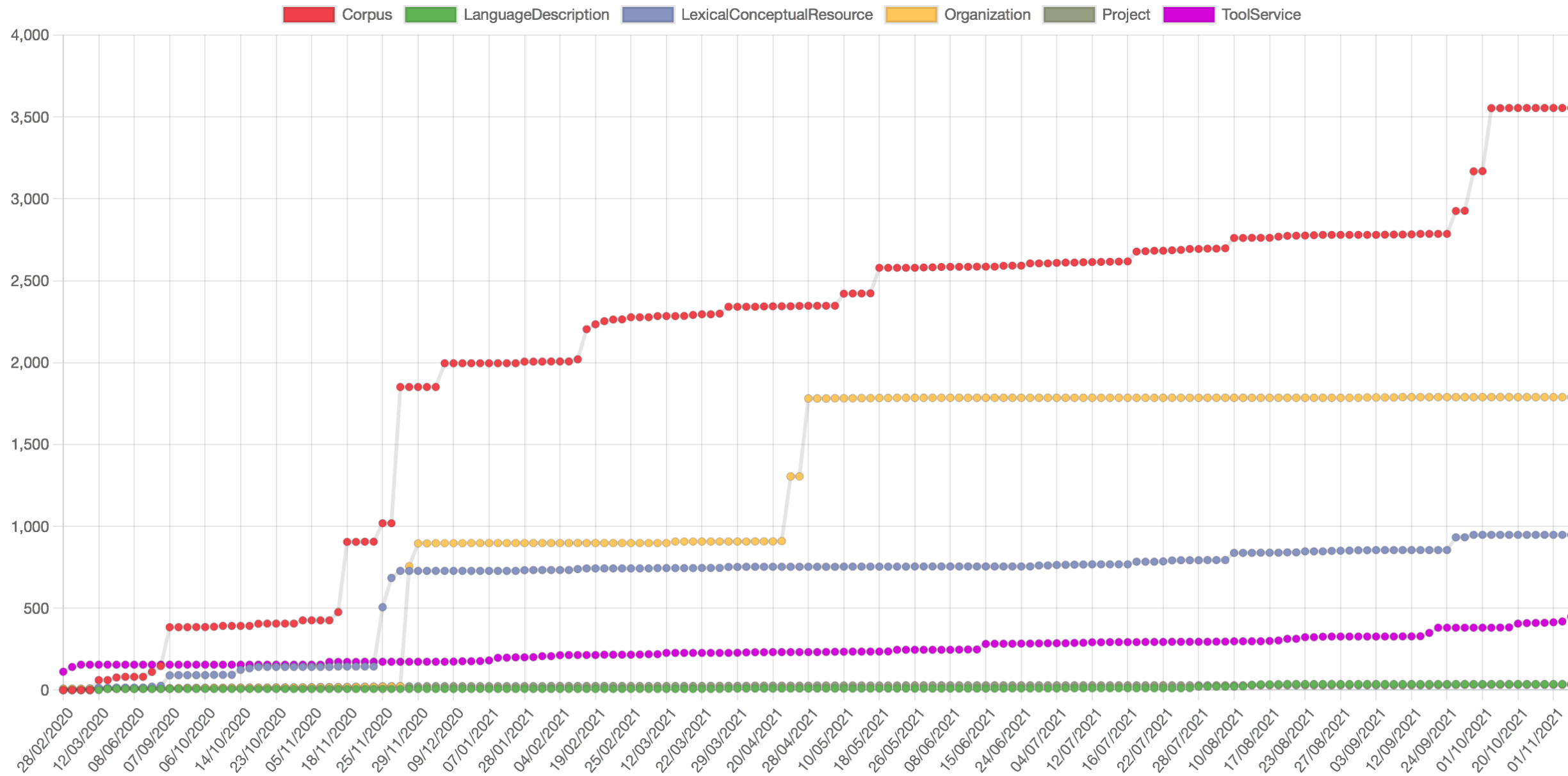
	01/12/2020 META-FORUM 2020	12/11/2021 META-FORUM 2021
Corpora	1851	3563
Tools/Services	173	450
Language Descriptions	7	37
Lexical Conceptual Resources	728	948
Projects	24	32
Total	3679	5030
Additional resources collected by ELE	–	approx. 6000
Total (online soon)	3679	11030



Number of Resources over Time



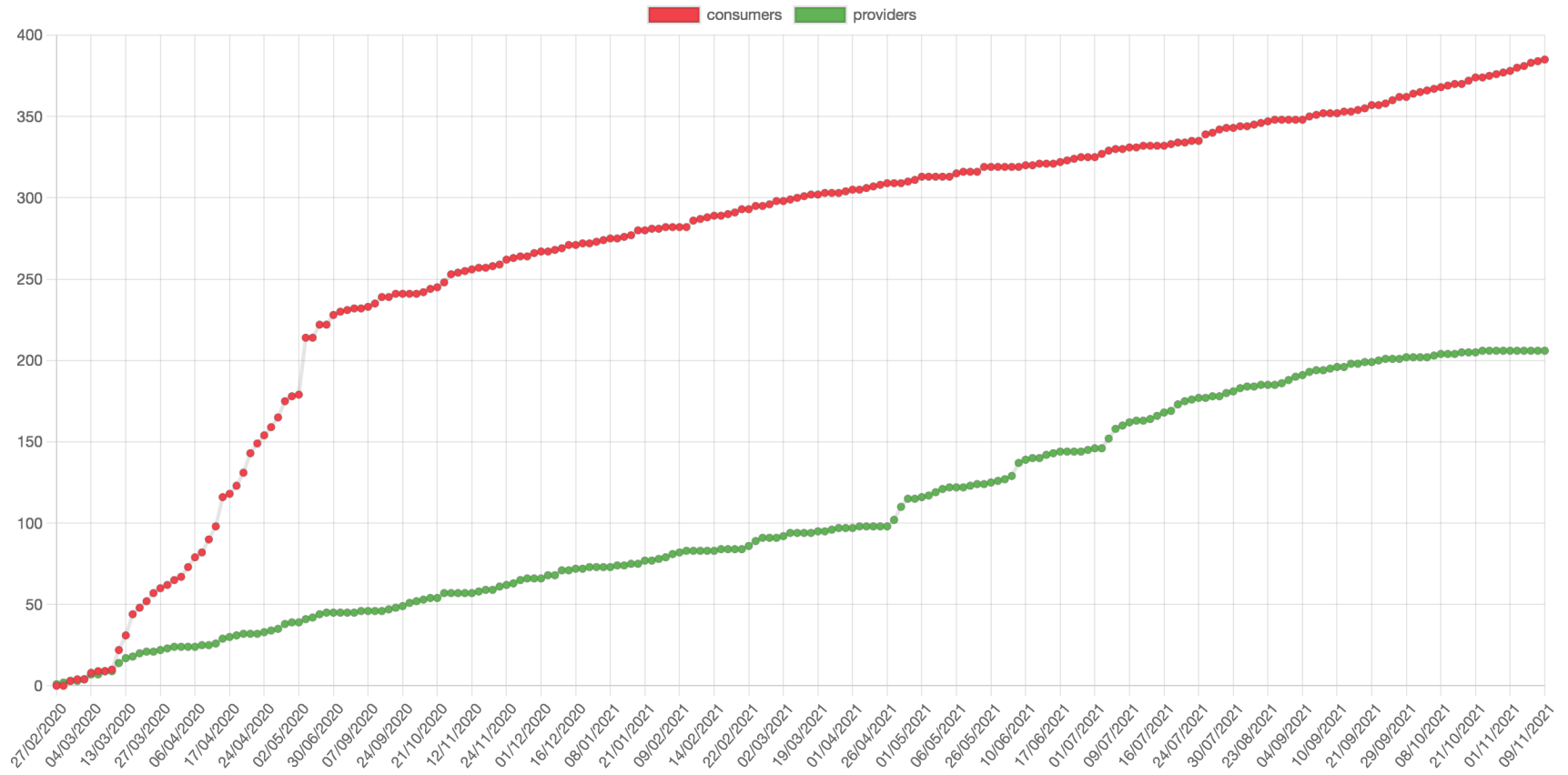
Number of Resources and Resource Types over Time



Number of Users



EUROPEAN LANGUAGE GRID

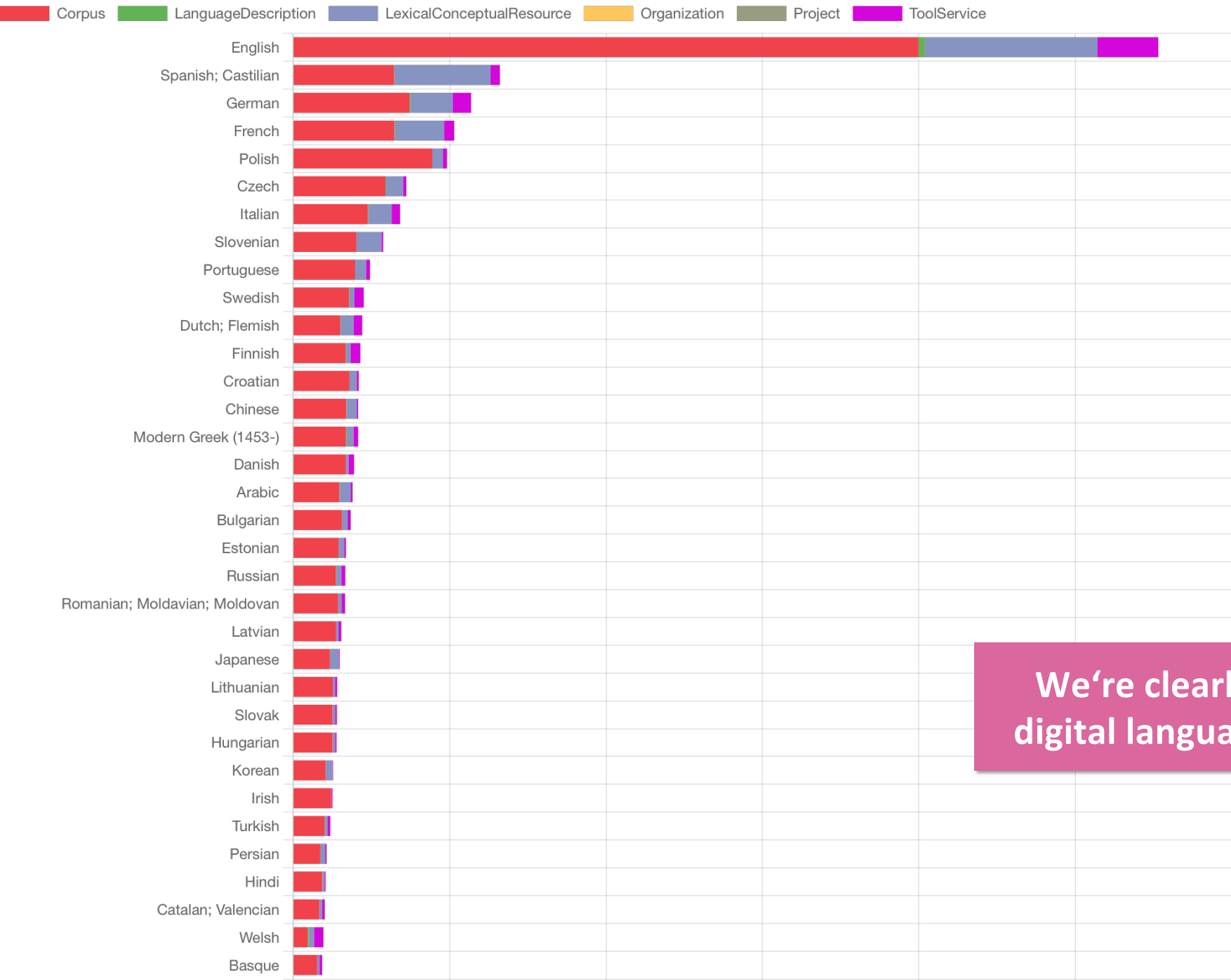


	01/12/2020 META-FORUM 2020	28/10/2021 META-FORUM 2021
Consumers	283	385
Providers	50	206
Total	333	591

Languages of the Resources



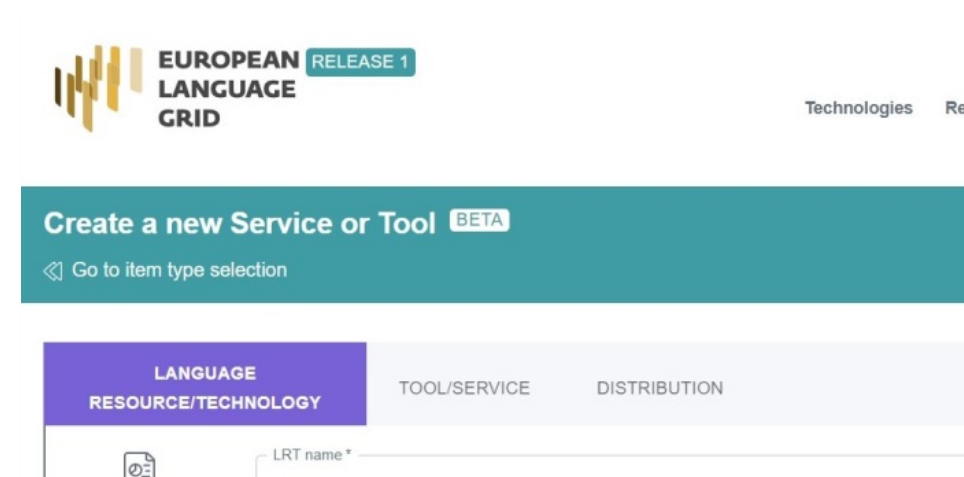
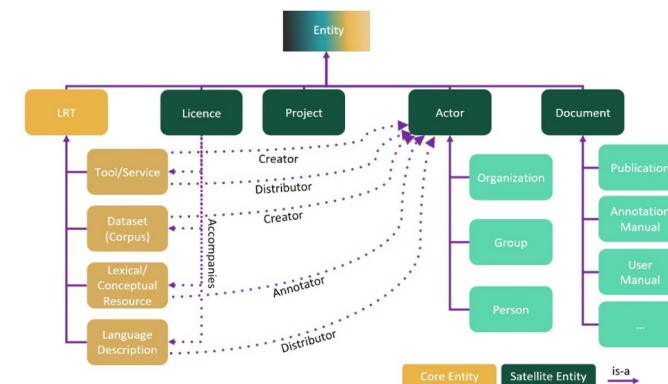
EUROPEAN LANGUAGE GRID



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

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



Example: Integration of the results of the project QURATOR into the ELG platform.

What does that mean?

- We created a page for the QURATOR project in the ELG catalogue
- We created for all QURATOR partners, too
- Two QURATOR services have already been made available through the ELG platform – their metadata records are associated with QURATOR
- More QURATOR resources will be made available in the future

Qurator
Curation Technologies

ART+COM

ada

SEM-TATION

ubermetrics

3pc
Neue Kommunikation

SBB

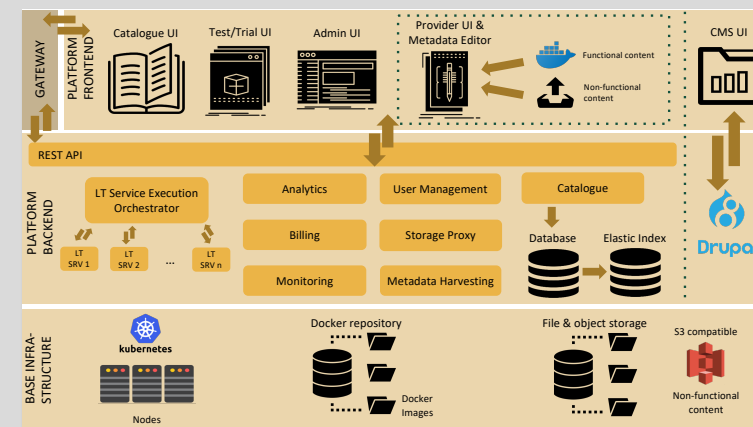
condat®

WIKIMEDIA
DEUTSCHLAND

DFK
Deutsches
Forschungszentrum
für Künstliche
Intelligenz GmbH

Fraunhofer
FOKUS

**EUROPEAN
LANGUAGE
GRID**



Go to catalogue

QURATOR – Curation Technologies. Flexible KI-Verfahren für die adaptive Analyse und kreative Generierung digitaler Inhalte in branchenübergreifenden Kontexten

Overview Related entities

In all domains and sectors, the demand for intelligent systems to support the processing and generation of digital content is rapidly increasing. The availability of vast amounts of content and the pressure to publish new content quickly and in rapid succession requires faster, more efficient and smarter processing and generation methods. With a consortium of ten partners from research and industry and a broad range of expertise in AI, Machine Learning and Language Technologies, the QURATOR project, funded by the G... [Read More](#)

- Keyword**
 - Language technology services
 - Multilingualism
 - Curation Technologies
- LT area**
 - Language Technology

Coordinator
 German Research Center for Artificial Intelligence
[Website](#)

- Participants**
- ART+COM AG [Website](#)
 - Condat AG [Website](#)
 - 3pc GmbH [Website](#)
 - Fraunhofer FOKUS [Website](#)
 - Semtation GmbH [Website](#)
 - Stiftung Preußischer Kulturbesitz [Website](#)
 - uberMetrics Technologies GmbH [Website](#)
 - Wikimedia Deutschland e.V. [Website](#)
 - Ada Health GmbH [Website](#)

Project information

[Website](#)

Project start date
2018-11-01

Project end date
2021-10-31

Funder

Federal Ministry of Education and Research
[Website](#)

Funding scheme category

IA

Funding country

Germany

Funding type

national funds

Grant number: Wachstumskern no.–03WKDA1A

Status
SIGNED

Related call
Wachstumskerne – Unternehmen Region

Related programme
Innovative regionale Wachstumskerne

Go to catalogue

QURATOR – Curation Technologies. Flexible KI-Verfahren für die adaptive Analyse und kreative Generierung digitaler Inhalte in branchenübergreifenden Kontexten

Overview Related entities

Funded Language Resources and Technologies

BERTNER German v1.0.0

BERTNER English v1.0.0

Similar integration of services or resources can be done by companies or research groups: all members of the European LT community can make their assets available via the ELG platform!

If you're interested in more details, please participate in the ELG tutorial tomorrow at META-FORUM 2021 (11:00-12:30).

BERTNER-DE

Version: v1.0.0

Functional service

Overview Download/Run Test/Try out Code samples

The Named Entity Recognition we are using two different approaches to recognise entities: (1) based on models and (2) based on dictionaries. This module has been implemented using the Apache OpenNLP tool. The model based approach is using the NameFinderME module that is provided with a specific model (language-dependent). We have trained separate models for each entity type we distinguish. This means that we look for persons, organisations and locations serially, with three corresponding models applied consecutively... [Read More](#)

Resource provider

German Research Center for Artificial Intelligence

[Website](#)

Funded by

QURATOR – Curation Technologies. Flexible KI-Verfahren für die adaptive Analyse und kreative Generierung digitaler Inhalte in branchenübergreifenden Kontexten

[Website](#)

Keyword
Named Entity Recognition
BERT
Language Models

Intended application
Named Entity Recognition

Input content resource

Language
German

Data format
JSON

Processing resource type
file

more

Function

Function
Named Entity Recognition

Language dependent
true

Output resource

Language
German

Data format
JSON

Processing resource type
file

more

Resource creator

German Research Center for Artificial Intelligence

[Website](#)

Publication date
2019-08-23

Evaluated: false

TRL: TRL3

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

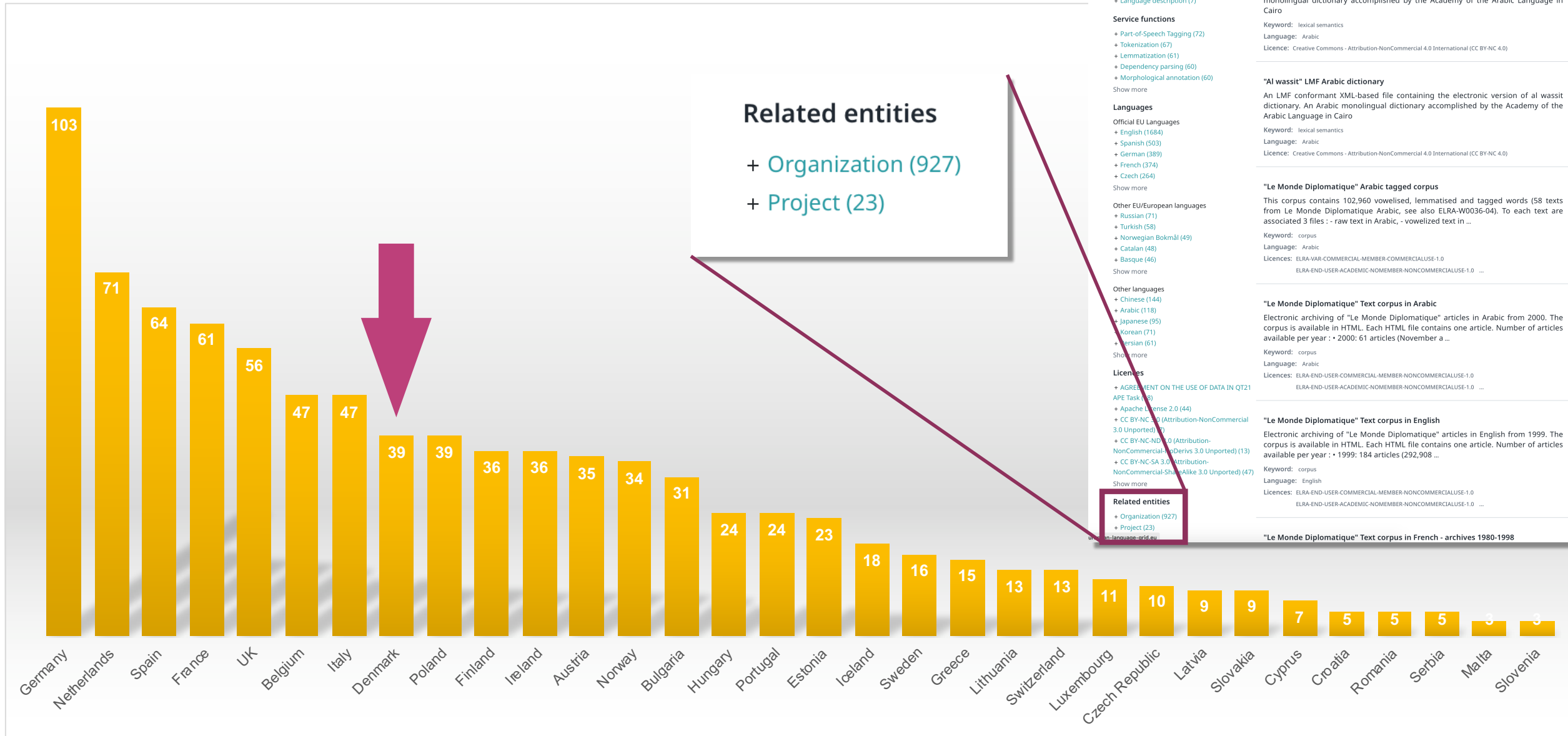
Other organisations (e.g., NGOs) that *purchase* or *use* Language Technologies

Funding agencies that support the development of Language Technologies



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

European LT Industry in the ELG



Related entities

- + Organization (927)
- + Project (23)

Clear all filters

- Language resources & technologies
 - + Corpus (1857)
 - + Lexical/Conceptual resource (728)
 - + Tool/Service (168)
 - + Language description (7)
- Service functions
 - + Part-of-Speech Tagging (72)
 - + Tokenization (67)
 - + Lemmatization (61)
 - + Dependency parsing (60)
 - + Morphological annotation (60)
- Languages
 - Official EU Languages
 - + English (1684)
 - + Spanish (503)
 - + German (389)
 - + French (374)
 - + Czech (264)
 - Other EU/European languages
 - + Russian (71)
 - + Turkish (58)
 - + Norwegian Bokmål (49)
 - + Catalan (48)
 - + Basque (46)
 - Other languages
 - + Chinese (144)
 - + Arabic (118)
 - + Japanese (95)
 - + Korean (71)
 - + Persian (61)
- Licences
 - + AGREEMENT ON THE USE OF DATA IN QT21 APE Task (8)
 - + Apache License 2.0 (44)
 - + CC BY-NC 3.0 (Attribution-NonCommercial 3.0 Unported) (7)
 - + CC BY-NC-ND 3.0 (Attribution-NonCommercial-NoDerivs 3.0 Unported) (13)
 - + CC BY-NC-SA 3.0 (Attribution-NonCommercial-ShareAlike 3.0 Unported) (47)
- Related entities**
 - + Organization (927)
 - + Project (23)

3710 search results

- "Al wassit" Arabic dictionary**

Lexical/Conceptual resource
114 views

An XML-based file containing the electronic version of al wassit dictionary. An Arabic monolingual dictionary accomplished by the Academy of the Arabic Language in Cairo

Keyword: lexical semantics
Language: Arabic
Licence: Creative Commons - Attribution-NonCommercial 4.0 International (CC BY-NC 4.0)
- "Al wassit" LMF Arabic dictionary**

Lexical/Conceptual resource
21 views

An LMF conformant XML-based file containing the electronic version of al wassit dictionary. An Arabic monolingual dictionary accomplished by the Academy of the Arabic Language in Cairo

Keyword: lexical semantics
Language: Arabic
Licence: Creative Commons - Attribution-NonCommercial 4.0 International (CC BY-NC 4.0)
- "Le Monde Diplomatique" Arabic tagged corpus**

Corpus
1 views

This corpus contains 102,960 vowelised, lemmatised and tagged words (58 texts from Le Monde Diplomatique Arabic, see also ELRA-W0036-04). To each text are associated 3 files : - raw text in Arabic, - vowelized text in ...

Keyword: corpus
Language: Arabic
Licences: ELRA-VAR-COMMERCIAL-MEMBER-COMMERCIALUSE-1.0 ... ELRA-END-USER-ACADEMIC-NOMEMBER-NONCOMMERCIALUSE-1.0 ...
- "Le Monde Diplomatique" Text corpus in Arabic**

Corpus
1 views

Electronic archiving of "Le Monde Diplomatique" articles in Arabic from 2000. The corpus is available in HTML. Each HTML file contains one article. Number of articles available per year : • 2000: 61 articles (November a ...

Keyword: corpus
Language: Arabic
Licences: ELRA-END-USER-COMMERCIAL-MEMBER-NONCOMMERCIALUSE-1.0 ... ELRA-END-USER-ACADEMIC-NOMEMBER-NONCOMMERCIALUSE-1.0 ...
- "Le Monde Diplomatique" Text corpus in English**

Corpus
1 views

Electronic archiving of "Le Monde Diplomatique" articles in English from 1999. The corpus is available in HTML. Each HTML file contains one article. Number of articles available per year : • 1999: 184 articles (292,908 ...

Keyword: corpus
Language: English
Licences: ELRA-END-USER-COMMERCIAL-MEMBER-NONCOMMERCIALUSE-1.0 ... ELRA-END-USER-ACADEMIC-NOMEMBER-NONCOMMERCIALUSE-1.0 ...
- "Le Monde Diplomatique" Text corpus in French - archives 1980-1998**

Corpus

Stakeholders and Users and Collaborators

ELG Open Calls: 15 pilot projects

Six ICT-29b research and innovation projects

ELE – European Language Equality

AI4EU – European AI on demand platform

Related projects: ELRC, Elexis, Lynx, MeMAD, Fandango, CEF INEA projects (MAPA, NTEU etc.) etc.

Core networks: META-NET, LT Innovate

Additional initiatives: CLAIRE, AI PPP, ECSPM, EFNIL, BDVA, NPLD, CLARIN, W3C, NGI, RDA, EOSC, OpenAIRE etc., ICT-48 projects



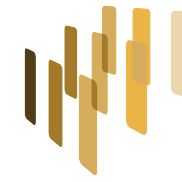
Global Under Resourced MEdia Translation



COMPRISE



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) (work in progress)

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

**EUROPEAN
LANGUAGE
EQUALITY**



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 TADIC

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

1. **Austria:** Dagmar Gromann, Zentrum für Translationswissenschaft, Universität Wien
2. **Belgium:** Walter Daelemans, Comp. Ling. and Psycholing. Res. Centre (CLiPS), University of Antwerp
3. **Bulgaria:** Svetla Koeva, Bulgarian Academy of Sciences
4. **Croatia:** Marko Tadic, Inst. of Linguistics, Faculty of Hum. and Social Science, University of Zagreb
5. **Cyprus:** Dora Loizidou, University of Cyprus
6.  **Czech Republic:** Jan Hajic, Inst. of Formal and Applied Linguistics, Charles University in Prague
7. **Denmark:** Bolette Sandford Pedersen, Centre for Lang. Technology, Department of Nordic Research, University of Copenhagen
8. **Estonia:** Kadri Vare, Department of Language, Estonian Ministry of Education and Research
9. **Finland:** Krister Linden, Department of Modern Languages, University of Helsinki
10. **France:** François Yvon, CNRS-LIMSI
11.  **Germany:** Georg Rehm, Speech and Language Technology Lab, DFKI
12.  **Greece:** Maria Gavriilidou, ILSP, R.C. "Athena"
13. **Hungary:** Tamás Várardi, Research Institute for Linguistics, Hungarian Academy of Sciences
14. **Iceland:** Eiríkur Rögnvaldsson, School of Humanities, University of Iceland
15. **Ireland:** Andy Way, ADAPT Centre and School of Computing, Dublin City University
16. **Italy:** Bernardo Magnini, Human Language Technology, Fondazione Bruno Kessler (FBK)
17. **Latvia:** Inguna Skadina, Institute of Mathematics and Computer Science, University of Latvia
18. **Lithuania:** Albina Aukšoriūtė, Institute of the Lithuanian Language
19. **Luxembourg:** Dimitra Anastasiou, Luxembourg Institute of Science and Technology
20. **Malta:** Michael Rosner, Department Intelligent Computer Systems, University of Malta
21. **Netherlands:** Jan Odijk, Utrecht Institute of Linguistics, Universiteit Utrecht
22. **Norway:** Kristine Eide, The Language Council of Norway – Språkrådet
23. **Poland:** Maciej Ogrodniczuk, Institute of Computer Science, Polish Academy of Sciences
24. **Portugal:** António Branco, Department of Informatics, University of Lisbon
25. **Romania:** Dan Tufis, Research Institute for Artificial Intelligence, Romanian Academy of Sciences
26. **Serbia:** Cvetana Krstev, Faculty of Mathematics, Belgrade University (UBG)
27. **Slovakia:** Radovan Garabik, Ludovít Stur Institute of Linguistics, Slovak Academy of Sciences
28. **Slovenia:** Simon Krek, Jozef Stefan Institute
29. **Spain:** Marta Villegas, Barcelona Supercomputing
30. **Sweden:** Jens Edlund, Speech, Music & Hearing/Språkbanken Tal, KTH Royal Institute of Technology
31. **Switzerland:** Hervé Boudlard, Idiap Research Institute
32.  **UK:** Kalina Bontcheva, Department of Computer Science, University of Sheffield



META-FORUM Conference Series

META-FORUM 2022 – June 08-10, Brussels, Belgium

Save the date!

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

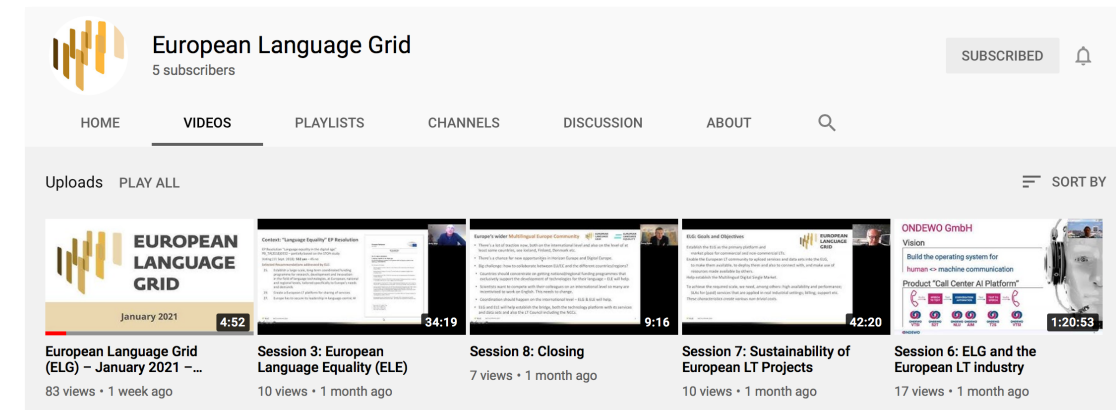
Solutions for Multilingual Europe

META-FORUM 2010 – November 17/18, Brussels, Belgium

Challenges for Multilingual Europe

The ELG YouTube channel provides recordings of all sessions of META-FORUM 2020 and also META-FORUM 2021 plus several other videos – *like and subscribe!*

<https://www.youtube.com/channel/UCarEHmsWT2JslcvvWkbhL4A>



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 shall

- 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

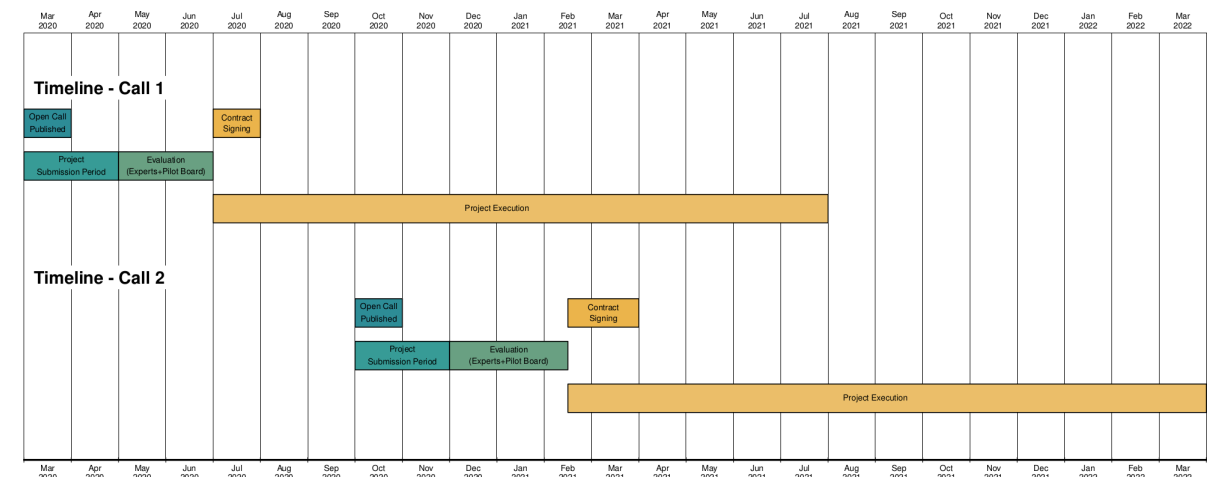
Total amount requested: 16.9M€

→ 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€
FZI Research Center for IT	Streaming Language Processing in Manufacturing ✓	DE	132,160€

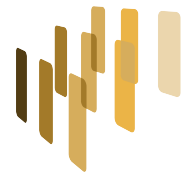
Open Call #2

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€

ELG's Sister Project European Language Equality (ELE)



**EUROPEAN
LANGUAGE
EQUALITY**



**EUROPEAN
LANGUAGE
GRID**

Consortium: 52 partners from all over Europe

Coordinator: ADAPT Centre (Dublin City University)

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

Runtime: 18 months – ELE & ELG will both finish in June 2022

Start on 1 January 2021

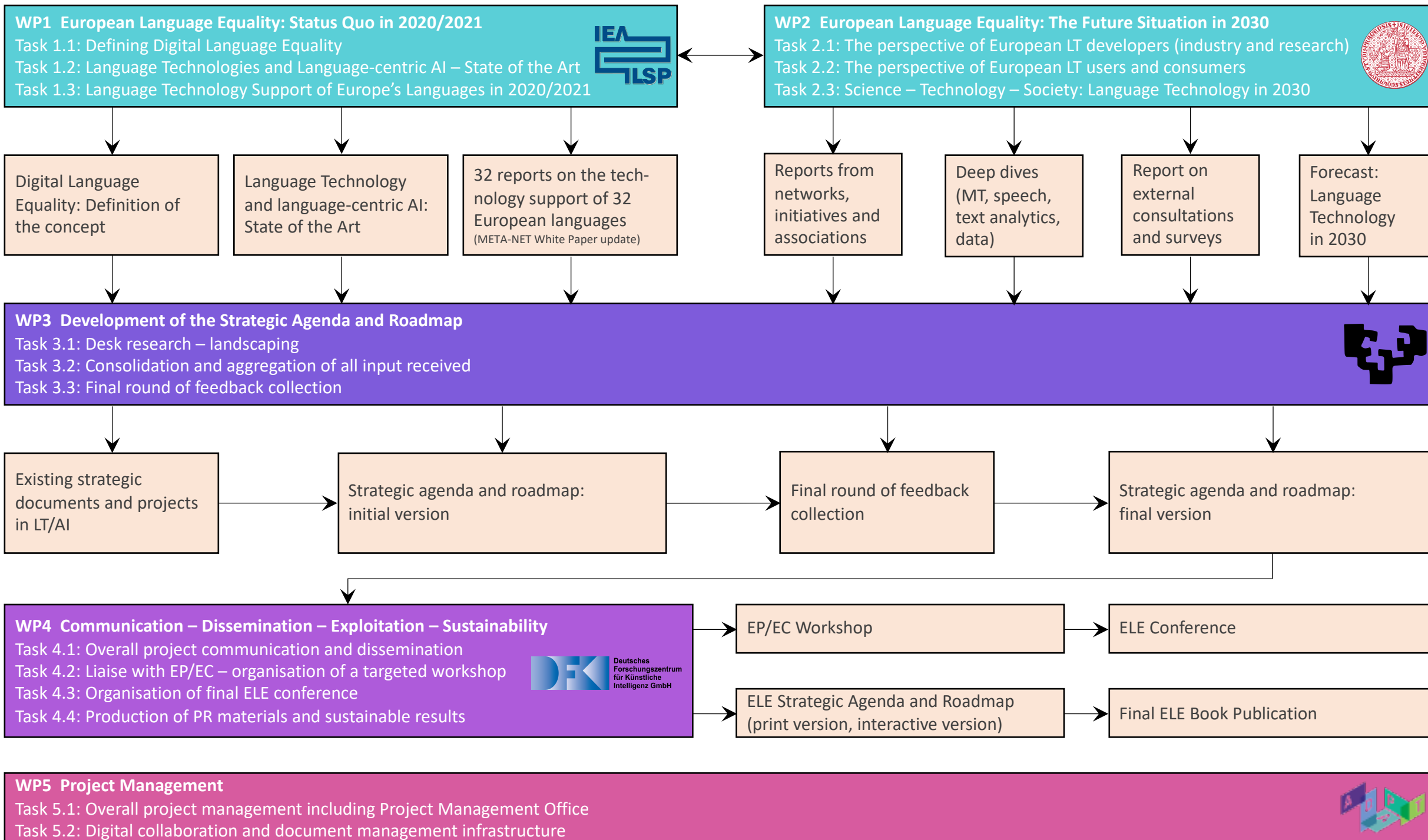
PP/PA (*not* Horizon 2020) – Pilot Project/Preparatory Action

Budget: 1,8M€ (Coordination and Support Action)

Nov. 2021: *more than halfway through ...*

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

**Work
Packages and
main
Deliverables**

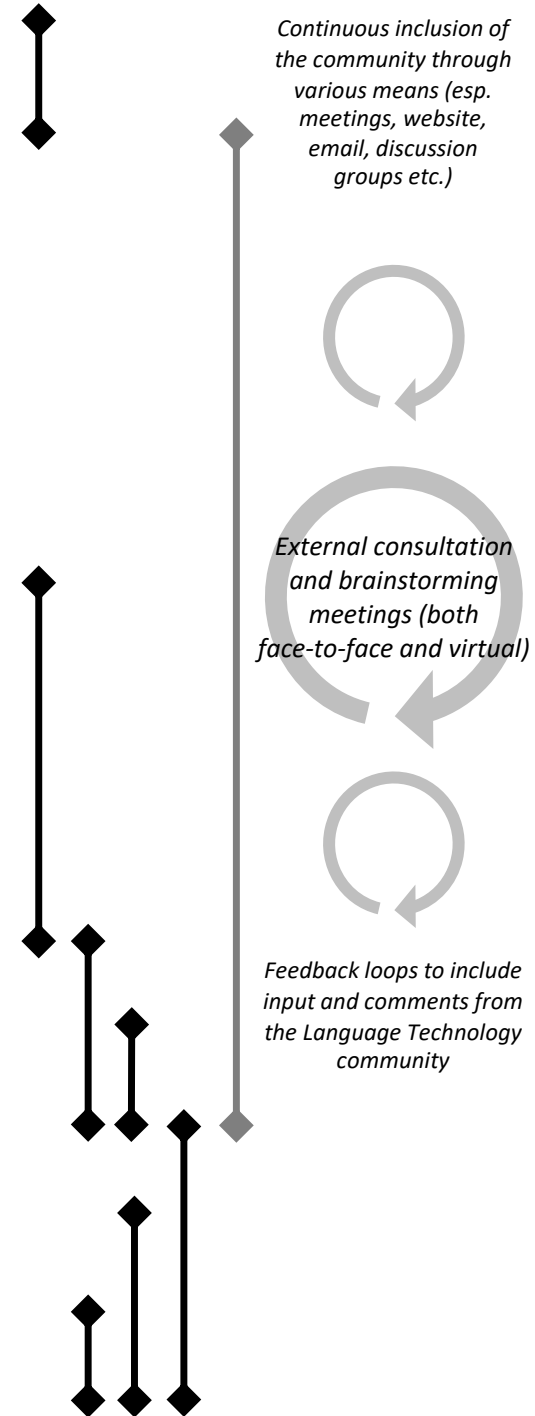


Digital collaboration and document management infrastructure (D5.1)

Project mgmt. report (D5.2)

Project mgmt. report (D5.3)
End of the ELE project

M1	ELE kick-off meeting
M2	
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	
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)
M18	Strategic agenda including roadmap – final version (D3.4) ELE EP/EC workshop (D4.3); ELE conference (D4.4); ELE book publication (D4.6)



Timeline

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

ELE

META FORUM 2021
Using the European Language Grid
15-17 November 2021
Virtual conference

Register for free at <https://meta-forum.eu>

Hosted by Berlin

If you're interested in more news regarding the ELE project, please join us tomorrow morning for the ELE session (09:00-10:30).

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

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

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:

<https://european-language-equality.eu>


Current State of Play in the Project

- **Digital Language Equality:** concept defined, dozens of *technological* and *contextual factors* longlisted, currently working on the selection of factors and the DLE formula
- **Surveys** conducted, now in the analysis phase – **Citizen survey** currently prepared for launch
- **Update of META-NET White Papers**, especially regarding the overview and qualitative analysis of technology support for Europe’s languages – plus: *dynamic and fully empirical dashboard*
- **Collection of LRs/LTs:** with the whole ELE consortium we collected 6000+ additional records
- **ELG:** these 6000+ records will be included in ELG, arriving at approx. *11,000 resources* in total
- **Dynamic DLE computation:** based on these 11,000 metadata records we can compute and visualise the DLE metric for all European languages *through a dedicated dashboard in ELG*
- **Continuously:** this approach enables us to monitor the development of DLE in Europe over time
- **ELG:** Making a new resource available through ELG will increase the DLE metric for that language
- Almost all analyses to be ready in approx. April 2022 – ELE book to be published in June 2022

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, which will operate and maintain the technology platform for the whole LT community as a joint marketplace.
- Contribute to **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**: validation of **ELG products**; attach more **data repositories**; include ELG in **relevant infrastructures** (e.g., NFDI, GAIA-X); **raise awareness** of LT for Europe; organise **META-FORUM 2022**; develop **ELG Release 3** (early 2022); establish the **ELG legal entity** (2022).



EUROPEAN LANGUAGE TECHNOLOGY
ELE ELG

European Language Technology

Fostering the European Language Technology community towards digital language equality in Europe by 2030!
Research · Brussels · 408 followers

[✓ Following](#) [Learn more](#) [More](#)

[Home](#) [About](#) [Posts](#) [Jobs](#) [People](#) [Videos](#)

About

This is the combined channel of the EU projects European Language Equality (ELE) and European Language Grid (ELG) – together working towards a joint network of language technology for Europe’s languages and digital language equality by 2030. Follow us for project updates, events and news from the European artificial...
[see more](#)

[See all details](#)



← **European Language Technology**
74 Tweets



EUROPEAN LANGUAGE TECHNOLOGY
ELE ELG

[Edit profile](#)

European Language Technology

@EuroLangTech

The channel of the EU projects European Language Equality and European Language Grid, fostering the LT community towards digital language equality by 2030!

📍 Europe [european-language-technology.eu](#) 📅 Joined June 2021

902 Following 303 Followers




Welcome to European Language Technology!
ELT

[Follow ELT on Twitter](#)
[Follow ELT on LinkedIn](#)

European Language Technology is the combined communication channel for the sister projects European Language Grid (ELG) and European Language Equality (ELE) – funded by the European Commission.

For further information about either project – their goals, consortium partners, and contact details etc. – please click below:

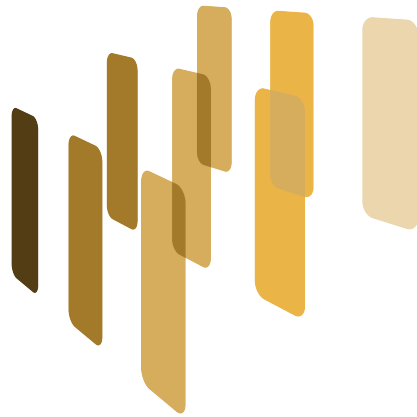
<https://www.linkedin.com/company/european-language-technology>

<https://twitter.com/EuroLangTech>

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

Subscribe to our newsletter (2500+ subscribers already)

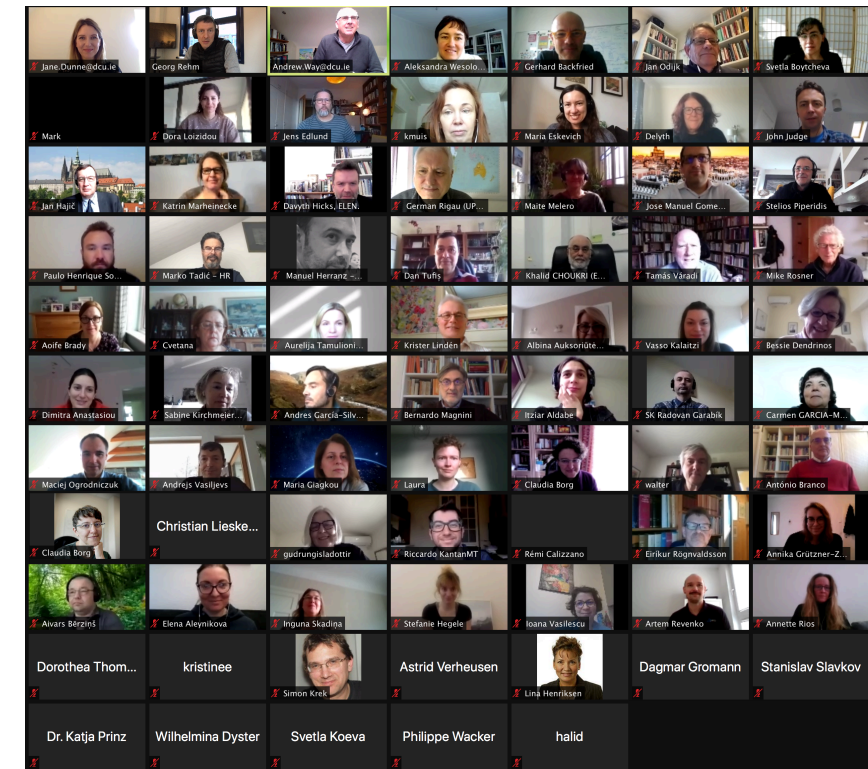
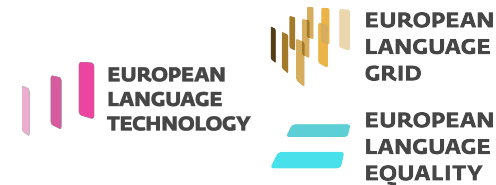




With many thanks to all colleagues in the ELG and ELE teams and consortia!



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).

Prof. Dr. Georg Rehm (DFKI GmbH, Germany)
Coordinator ELG, Co-coordinator ELE

16-11-2021 ELG NCC Workshop Denmark – Danish LT Conference 2021
<http://www.european-language-grid.eu>
<http://www.european-language-equality.eu>