The European Cloud Platform for Language Technology, Natural Language Processing and Language-Centric AI

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

18-10-2021 ELG NCC Workshop Czech Republic & Slovakia
http://www.european-language-grid.eu
http://www.european-language-equality.eu
European Language Grid

- History and context
- European Language Grid
- European Language Equality
- Summary and next steps
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
<table>
<thead>
<tr>
<th>Year</th>
<th>Event/Initiative</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010</td>
<td><strong>META FORUM 2010</strong></td>
<td>“Challenges for Multilingual Europe” (November 17/18, 2010)</td>
</tr>
<tr>
<td>2011</td>
<td><strong>META FORUM 2011</strong></td>
<td>Conference “A Strategy for Multilingual Europe” (Brussels – June 20/21, 2011)</td>
</tr>
<tr>
<td>2012</td>
<td><strong>META FORUM 2012</strong></td>
<td>Conference “A Strategy for Multilingual Europe” (Brussels – June 20/21, 2012)</td>
</tr>
<tr>
<td>2016</td>
<td><strong>META FORUM 2016</strong></td>
<td>Conference “Beyond Multilingual Europe” (Lisbon – July 4/5, 2016) Strategic Research and Innovation Agenda titled “Language as a Data Type and Key Challenge for Big Data” (July 2016)</td>
</tr>
<tr>
<td>2017</td>
<td><strong>META FORUM 2017</strong></td>
<td>Launch of the Strategic Research and Innovation Agenda titled “Language Technologies for Multilingual Europe – Towards a Human Language Project” (December 2017)</td>
</tr>
<tr>
<td>2018</td>
<td><strong>META FORUM 2018</strong></td>
<td>ELG Final Proposal Submission Final submission on Feb. 20, 2018</td>
</tr>
</tbody>
</table>

**META** - European Language Grid

**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-NET4U** - EU-funded project (ICT Policy Support Programme) to enhance the European Linguistic Infrastructure (2011 – 2013)

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

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

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

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

* Official EU language
<table>
<thead>
<tr>
<th>Excellent support</th>
<th>Good support</th>
<th>Moderate support</th>
<th>Fragmentary support</th>
<th>Weak or no support through LT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Czech, Dutch, French, German, Italian, Portuguese, Spanish</td>
<td>Czech, Dutch, Finnish, French, German, Hungarian, Italian, Polish, Swedish</td>
<td>English</td>
<td>Catalan, Dutch, German, Hungarian, Italian, Polish, Romanian</td>
<td>Basque, Bulgarian, Catalan, Croatian, Czech, Danish, Estonian, Finnish, Galician, Greek, Icelandic, Irish, Latvian, Lithuanian, Maltese, Norwegian, Portuguese, Slovak, Slovene, Swedish, Welsh</td>
</tr>
<tr>
<td>Czech, Dutch, Finnish, French, German, Hungarian, Italian, Polish, Spanish</td>
<td>Dutch, French, German, Italian, Spanish</td>
<td>English</td>
<td>Basque, Bulgarian, Catalan, Danish, Estonian, Finnish, Galician, Greek, Hungarian, Norwegian, Polish, Portuguese, Romanian, Slovak, Slovene, Swedish</td>
<td>Croatian, Estonian, Icelandic, Irish, Latvian, Lithuanian, Maltese, Serbian, Welsh</td>
</tr>
<tr>
<td>English</td>
<td>French, Spanish</td>
<td>English</td>
<td>Basque, Bulgarian, Catalan, Czech, Danish, Finnish, Galician, Greek, Hungarian, Norwegian, Polish, Portuguese, Romanian, Slovak, Slovene, Swedish</td>
<td>Croatian, Icelandic, Latvian, Lithuanian, Maltese, Romanian, Welsh</td>
</tr>
<tr>
<td>English</td>
<td>Czech, Dutch, Finnish, French, German, Italian, Portuguese, Spanish</td>
<td>English</td>
<td>Basque, Bulgarian, Catalan, Danish, Estonian, Finnish, Galician, Greek, Hungarian, Irish, Norwegian, Polish, Serbian, Slovak, Slovene, Swedish</td>
<td>Icelandic, Irish, Latvian, Lithuanian, Maltese, Welsh</td>
</tr>
</tbody>
</table>
Languages with names in red have little or no MT support.

We carried out the study in 2011/2012. While support for many languages, and also overall, has improved in the meantime thanks to Deep Learning, the bigger picture appears to remain mostly the same.

<table>
<thead>
<tr>
<th>Language</th>
<th>Level of support</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>Excellent</td>
</tr>
<tr>
<td>French</td>
<td>Good</td>
</tr>
<tr>
<td>Spanish</td>
<td>Moderate</td>
</tr>
<tr>
<td>Dutch</td>
<td>Fragmentary</td>
</tr>
<tr>
<td>German</td>
<td>Weak/none</td>
</tr>
<tr>
<td>Italian</td>
<td>Poor</td>
</tr>
<tr>
<td>Czech</td>
<td>Poor</td>
</tr>
<tr>
<td>Hungarian</td>
<td>Poor</td>
</tr>
<tr>
<td>Polish</td>
<td>Poor</td>
</tr>
<tr>
<td>Catalan</td>
<td>Poor</td>
</tr>
<tr>
<td>Finnish</td>
<td>Poor</td>
</tr>
<tr>
<td>Portuguese</td>
<td>Poor</td>
</tr>
<tr>
<td>Swedish</td>
<td>Poor</td>
</tr>
<tr>
<td>Basque</td>
<td>Poor</td>
</tr>
<tr>
<td>Bulgarian</td>
<td>Poor</td>
</tr>
<tr>
<td>Danish</td>
<td>Poor</td>
</tr>
<tr>
<td>Galician</td>
<td>Poor</td>
</tr>
<tr>
<td>Greek</td>
<td>Poor</td>
</tr>
<tr>
<td>Norwegian</td>
<td>Poor</td>
</tr>
<tr>
<td>Romanian</td>
<td>Poor</td>
</tr>
<tr>
<td>Slovak</td>
<td>Poor</td>
</tr>
<tr>
<td>Slovene</td>
<td>Poor</td>
</tr>
<tr>
<td>Estonian</td>
<td>Poor</td>
</tr>
<tr>
<td>Serbian</td>
<td>Poor</td>
</tr>
<tr>
<td>Croatian</td>
<td>Poor</td>
</tr>
<tr>
<td>Irish</td>
<td>Poor</td>
</tr>
<tr>
<td>Icelandic</td>
<td>Poor</td>
</tr>
<tr>
<td>Latvian</td>
<td>Poor</td>
</tr>
<tr>
<td>Lithuanian</td>
<td>Poor</td>
</tr>
<tr>
<td>Maltese</td>
<td>Poor</td>
</tr>
<tr>
<td>Welsh</td>
<td>Poor</td>
</tr>
</tbody>
</table>


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
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.
European Language Grid

Grid Community
- NCCs set up
- Communication Plan developed
- First ELG Conference
- LT Board Meeting
- Second ELG Conference
- LT Board Meeting
- Pilots: First Open Call
- Pilots: Second Open Call
- Third ELG Conference
- LT Board Meeting

Grid Content
- Test Services set up
- Services, Tools, Components: First Release
- Services, Tools, Components: Second Release
- Services, Tools, Components: Final Release
- Data Sets: First Release
- Data Sets: Second Release
- Data Sets: Final Release

Grid Platform
- Cloud Platform set up
- First Demo
- Metadata Schema
- Platform: First Release
- Platform: Second Release
- Platform: Final Release

Mgmt.
- Project Website
- Corporate Identity
- Sustainability Plan (First Version)
- Sustainability Plan (Final Version)

Corporate Identity
ELG Legal Entity
European Language Grid – Current State of Play (September 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 (Oct. 2021):**

- 428 functional services and tools
- 3500 corpora and data sets
- 908 lexical/conceptual resources
- 25 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 2 made available at the end of March 2021**

**ELG Release 3 to be made available in early 2022.**
European Language Grid – Quick Demo

- 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://www.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.

https://live.european-language-grid.eu
Towards the Primary Platform for Language Technologies in Europe

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

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

Community
Organizations, projects, events etc. in European Language technology field
48 search results

2007 CoNLL Shared Task - Basque, Catalan, Czech & Turkish
2007 CoNLL Shared Task - Basque, Catalan, Czech & Turkish consists of dependency treebanks in four languages used as part of the CoNLL 2007 shared task on multi-lingual dependency parsing and domain adaptation. The langu ...
Keyword: corpus
Languages: Basque Turkish Czech
Licenses: ELIA-END-USER-ACADEMIC-MEMBER-NONCOMMERCIALE-1.0
ELIA-END-USER-COMMERCIAL-MEMBER-NONCOMMERCIALE-1.0 ...

Annotated corpora and tools of the PARSEME Shared Task on Automatic Identification of Verbal Multiword Expressions (edition 1.1)
This multilingual resource contains corpora in which verbal MWEs have been manually annotated. VMWES include idioms (let the cat out of the bag), light-verb constructions (make a decision), verb-particle constructions (g... Keywords: Multiword expressions verbal multiword expressions light verb constructions verb-particle constructions Languages: English Modern Greek Bulgarian German Polish ...
Licence: PARSEME Shared Task Data v.1.1 Agreement

Annotated corpora and tools of the PARSEME Shared Task on Semi-Supervised Identification of Verbal Multiword Expressions (edition 1.2)
This multilingual resource contains corpora in which verbal MWEs have been manually annotated, gathered at the occasion of the 1.2 edition of the PARSEME Shared Task on semi-supervised Identification of Verbal MWEs (2020 ... Keywords: multiword expressions verbal multiword expressions light verb construction
85 search results for uditpipe basque

UDPipe Basque: Morphiysontactic Analysis of Raw Text
UDPipe performs segmentation, tokenization, POS tagging, morphological analysis, lemmatization and dependency parsing of raw Basque texts.
Keywords: Tokenization Part of speech tagging Lemmatisation Morphological analyser
Language: Basque
Licences: Mozilla Public License 2.0 Creative Commons Attribution Non Commercial Share Alike 4.0 International

Basque WordNet
The Basque WordNet is a lexical database including information about Basque words. It is an extension of WordNet 1.6, a lexical database for English developed at the Princeton University. The Basque WordNet is tightly al...
Keywords: lexicocentricresource
Language: Basque
Licences: ELRA-END-USER-COMMERCIAL-MEMBER-NONCOMMERCIALUSE-1.0 ELRA-END-USER-ACADEMIC-NOMEMBER-NONCOMMERCIALUSE-1.0 ...

Elhuyar Basque ASR
Automatic Speech Recognition service for Basque
Keywords: automatic speech recognition Basque
Language: Basque
Licence: Apache License 2.0

Show more

Languages

Language resources & technologies

Service functions

Clear all filters

Show more
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 Python SDK – Overview

- Brings main ELG functionalities to Python.
- Integrate ELG services into pipelines.
- Repository available on Gitlab: https://gitlab.com/european-language-grid/platform/python-client

Currently available functionalities:
- Search the ELG catalogue
- Download ELG resources
- Call ELG services
- Compare multiple ELG services on various inputs (benchmarking)

Coming up:
- Contribute resources/services to ELG using Python through automated deployment
- Enable running ELG services locally (by running Docker images locally)
- Improve search experience
- Many additional features …
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
European LT Industry in the ELG

Related entities
+ Organization (927)
+ Project (23)
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
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.*
<table>
<thead>
<tr>
<th>Country</th>
<th>Name</th>
<th>Institution/Department</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austria</td>
<td>Dagmar Gromann</td>
<td>Zentrum für Translationswissenschaft, Universität Wien</td>
</tr>
<tr>
<td>Belgium</td>
<td>Walter Daelemans</td>
<td>Comp. Ling. and Psycholing. Res. Centre (CLiPS), University of Antwerp</td>
</tr>
<tr>
<td>Bulgaria</td>
<td>Svetla Koeva</td>
<td>Bulgarian Academy of Sciences</td>
</tr>
<tr>
<td>Croatia</td>
<td>Marko Tadic</td>
<td>Inst. of Linguistics, Faculty of Hum. and Social Science, University of Zagreb</td>
</tr>
<tr>
<td>Cyprus</td>
<td>Dora Loizidou</td>
<td>University of Cyprus</td>
</tr>
<tr>
<td>Czech Republic</td>
<td>Jan Hajic</td>
<td>Inst. of Formal and Applied Linguistics, Charles University in Prague</td>
</tr>
<tr>
<td>Denmark</td>
<td>Bolette Sandford Pedersen</td>
<td>Centre for Lang. Technology, Department of Nordic Research, University of Copenhagen</td>
</tr>
<tr>
<td>Estonia</td>
<td>Kadri Vare</td>
<td>Department of Language, Estonian Ministry of Education and Research</td>
</tr>
<tr>
<td>Finland</td>
<td>Krister Linden</td>
<td>Department of Modern Languages, University of Helsinki</td>
</tr>
<tr>
<td>France</td>
<td>François Yvon</td>
<td>CNRS-LIMSI</td>
</tr>
<tr>
<td>Germany</td>
<td>Georg Rehm</td>
<td>Speech and Language Technology Lab, DFKI</td>
</tr>
<tr>
<td>Greece</td>
<td>Maria Gavrilidou</td>
<td>ILSP, R.C. “Athena”</td>
</tr>
<tr>
<td>Hungary</td>
<td>Tamás Várárdi</td>
<td>Research Institute for Linguistics, Hungarian Academy of Sciences</td>
</tr>
<tr>
<td>Iceland</td>
<td>Eirikur Rögnvaldsson</td>
<td>School of Humanities, University of Iceland</td>
</tr>
<tr>
<td>Ireland</td>
<td>Andy Way</td>
<td>ADAPT Centre and School of Computing, Dublin City University</td>
</tr>
<tr>
<td>Italy</td>
<td>Bernardo Magnini</td>
<td>Human Language Technology, Fondazione Bruno Kessler (FBK)</td>
</tr>
<tr>
<td>Latvia</td>
<td>Inguna Skadina</td>
<td>Institute of Mathematics and Computer Science, University of Latvia</td>
</tr>
<tr>
<td>Lithuania</td>
<td>Albina Auksořičtě</td>
<td>Institute of the Lithuanian Language</td>
</tr>
<tr>
<td>Luxembourg</td>
<td>Dimitra Anastasiou</td>
<td>Luxembourg Institute of Science and Technology</td>
</tr>
<tr>
<td>Malta</td>
<td>Michael Rosner</td>
<td>Department Intelligent Computer Systems, University of Malta</td>
</tr>
<tr>
<td>Netherlands</td>
<td>Jan Odijk</td>
<td>Utrecht Institute of Linguistics, Universiteit Utrecht</td>
</tr>
<tr>
<td>Norway</td>
<td>Kristine Eide</td>
<td>The Language Council of Norway – Språkrådet</td>
</tr>
<tr>
<td>Poland</td>
<td>Maciej Ogrodniczuk</td>
<td>Institute of Computer Science, Polish Academy of Sciences</td>
</tr>
<tr>
<td>Portugal</td>
<td>António Branco</td>
<td>Department of Informatics, University of Lisbon</td>
</tr>
<tr>
<td>Romania</td>
<td>Dan Tufis</td>
<td>Research Institute for Artificial Intelligence, Romanian Academy of Sciences</td>
</tr>
<tr>
<td>Slovakia</td>
<td>Radovan Garabik</td>
<td>Ludovit Stur Institute of Linguistics, Slovak Academy of Sciences</td>
</tr>
<tr>
<td>Slovenia</td>
<td>Simon Krek</td>
<td>Jozef Stefan Institute</td>
</tr>
<tr>
<td>Spain</td>
<td>Marta Villegas</td>
<td>Barcelona Supercomputing</td>
</tr>
<tr>
<td>Sweden</td>
<td>Jens Edleng</td>
<td>Speech, Music &amp; Hearing/Språkbanken Tal, KTH Royal Institute of Technology</td>
</tr>
<tr>
<td>Switzerland</td>
<td>Hervé Bourlard</td>
<td>Idiap Research Institute</td>
</tr>
<tr>
<td>UK</td>
<td>Kalina Bontcheva</td>
<td>Department of Computer Science, University of Sheffield</td>
</tr>
</tbody>
</table>
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 – Save the date!

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

All META-FORUM 2020 sessions and a five-minute screencast video explaining the current state of play of the ELG platform are available on the European Language Grid YouTube channel:

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

• We’ve identified several ways and approaches of covering the costs on a long-term basis.

• Establish consensus for a sustainable operational model.

• Options: a) for-profit or b) not-for-profit company, c) association, d) foundation

• Business Model Canvas prepared

• First set of approx. 20 products defined

• Next up: market validation of the product ideas

• Q1/Q2 2022: Establish legal entity
ELG’s Sister Project European Language Equality (ELE)

Consortium: 52 partners from all over Europe
Coordinator: ADAPT Centre (Dublin City University)
Co-Coordinator: DFKI
Objective: development of a strategic research, innovation and deployment agenda to achieve digital language equality in Europe by 2030
Runtime: 18 months – ELE and ELG will both finish up in June 2022 with a joint META-FORUM 2022
Started on 1 January 2021

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

We invite all language communities, all LT developing companies and all LT/NLP/Speech research teams to participate! Please participate in our two surveys – one for LT developers, one for LT users (URLs on the ELE website).
**Work Packages and main Deliverables**

**WP1 European Language Equality: Status Quo in 2020/2021**
- Task 1.1: Defining Digital Language Equality
- Task 1.2: Language Technologies and Language-centric AI – State of the Art
- Task 1.3: Language Technology Support of Europe’s Languages in 2020/2021

**WP2 European Language Equality: The Future Situation in 2030**
- Task 2.1: The perspective of European LT developers (industry and research)
- Task 2.2: The perspective of European LT users and consumers
- Task 2.3: Science – Technology – Society: Language Technology in 2030

**WP3 Development of the Strategic Agenda and Roadmap**
- Task 3.1: Desk research – landscaping
- Task 3.2: Consolidation and aggregation of all input received
- Task 3.3: Final round of feedback collection

**WP4 Communication – Dissemination – Exploitation – Sustainability**
- Task 4.1: Overall project communication and dissemination
- Task 4.2: Liaise with EP/EC – organisation of a targeted workshop
- Task 4.3: Organisation of final ELE conference
- Task 4.4: Production of PR materials and sustainable results

**WP5 Project Management**
- Task 5.1: Overall project management including Project Management Office
- Task 5.2: Digital collaboration and document management infrastructure
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.

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
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 2021 and 2022; develop ELG Release 3 (early 2022); establish the ELG legal entity (late 2021).
The European Language Technology Landscape in 2020: Language-Centric and Human-Centric AI for Cross-Cultural Communication to Multilingual Europe


Abstract

European Knowledge and Information Society (IKIS) and the European Research Area (ERA) in the 21st century are characterized by an exponential increase of digitalized information. Different types of data, new applications and services are transforming our societies. Knowledge communication and exchange are essential to support research and innovation activities. The European Research Council (ERC) and Horizon Europe have developed a new funding programme in 2020, Horizon Europe (HE), to consolidate and support European science and research. HE aims to boost the European Research Area for the benefit of society and Europe. The HE research landscape is highly diverse and involves many stakeholders, including researchers, entrepreneurs, SMEs, industries, funding agencies, citizens and citizens. The goal of HE is to create new synergies and to foster knowledge exchange and innovation.

1. Introduction

The European Research Council (ERC) and Horizon Europe (HE) aim to promote and support research in Europe. HE is the new European Research Framework Programme (EF) in the period 2021-2027. It is based on a ten-year Horizon Programme, which has a unique focus on excellence, ambition, robustness, and impact. HE focuses on three key pillars: excellence, impact, and sustainability. The HE research landscape is highly diverse and involves many stakeholders, including researchers, entrepreneurs, SMEs, industries, funding agencies, citizens and citizens. The goal of HE is to create new synergies and to foster knowledge exchange and innovation.

2. Objectives

The HE research landscape is highly diverse and involves many stakeholders, including researchers, entrepreneurs, SMEs, industries, funding agencies, citizens and citizens. The goal of HE is to create new synergies and to foster knowledge exchange and innovation.

2.1. Horizon Europe vs. Horizon 2020

HE focuses on excellence, impact, and sustainability. The HE research landscape is highly diverse and involves many stakeholders, including researchers, entrepreneurs, SMEs, industries, funding agencies, citizens and citizens. The goal of HE is to create new synergies and to foster knowledge exchange and innovation.

2.2. Objectives

HE focuses on excellence, impact, and sustainability. The HE research landscape is highly diverse and involves many stakeholders, including researchers, entrepreneurs, SMEs, industries, funding agencies, citizens and citizens. The goal of HE is to create new synergies and to foster knowledge exchange and innovation.

3. Horizon Europe: a new Framework Programme

HE is the new European Research Framework Programme (EF) in the period 2021-2027. It is based on a ten-year Horizon Programme, which has a unique focus on excellence, ambition, robustness, and impact. HE focuses on three key pillars: excellence, impact, and sustainability. The HE research landscape is highly diverse and involves many stakeholders, including researchers, entrepreneurs, SMEs, industries, funding agencies, citizens and citizens. The goal of HE is to create new synergies and to foster knowledge exchange and innovation.

4. Horizon Europe: the new Framework Programme

HE is the new European Research Framework Programme (EF) in the period 2021-2027. It is based on a ten-year Horizon Programme, which has a unique focus on excellence, ambition, robustness, and impact. HE focuses on three key pillars: excellence, impact, and sustainability. The HE research landscape is highly diverse and involves many stakeholders, including researchers, entrepreneurs, SMEs, industries, funding agencies, citizens and citizens. The goal of HE is to create new synergies and to foster knowledge exchange and innovation.

5. Horizon Europe: the new Framework Programme

HE is the new European Research Framework Programme (EF) in the period 2021-2027. It is based on a ten-year Horizon Programme, which has a unique focus on excellence, ambition, robustness, and impact. HE focuses on three key pillars: excellence, impact, and sustainability. The HE research landscape is highly diverse and involves many stakeholders, including researchers, entrepreneurs, SMEs, industries, funding agencies, citizens and citizens. The goal of HE is to create new synergies and to foster knowledge exchange and innovation.

6. Horizon Europe: the new Framework Programme

HE is the new European Research Framework Programme (EF) in the period 2021-2027. It is based on a ten-year Horizon Programme, which has a unique focus on excellence, ambition, robustness, and impact. HE focuses on three key pillars: excellence, impact, and sustainability. The HE research landscape is highly diverse and involves many stakeholders, including researchers, entrepreneurs, SMEs, industries, funding agencies, citizens and citizens. The goal of HE is to create new synergies and to foster knowledge exchange and innovation.
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

https://www.linkedin.com/company/european-language-technology
https://twitter.com/EuroLangTech
https://www.european-language-technology.eu
The European Language Grid has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement № 825627 (ELG).

Thank you!

Prof. Dr. Georg Rehm (DFKI GmbH, Germany)
Coordinator ELG, Co-coordinator ELE
18-10-2021 ELG NCC Workshop Czech Republic & Slovakia
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!