Proceedings of the Language Technologies for All (LT4All) , pages 131–136
Paris, UNESCO Headquarters, 5-6 December, 2019.

European Language Grid: Language Technologies for Europe

Georg Rehm
DFKI GmbH, Germany
georg.rehm@dfki.de

Abstract

With 24 official EU and many additional languages, multilingualism in Europe and an inclusive Digital Single Market can only be enabled through Language Technologies (LTs). European LT business is dominated by hundreds of SMEs and a few large players. Many are world-class, with technologies that outperform the global players. However, European LT business is also fragmented – by nation states, languages, verticals and sectors, significantly holding back its impact. The European Language Grid (ELG) project addresses this fragmentation by establishing the ELG as the primary platform for LT in Europe. The ELG is a scalable cloud platform, providing, in an easy-to-integrate way, access to hundreds of commercial and non-commercial LTs for all European languages. The ELG will enable the commercial and non-commercial European LT community to deposit and upload their technologies and data sets, to deploy them through the ELG, and to connect with other resources.

Keywords: LR Infrastructures and Architectures, LR National/International Projects, Tools, Systems, Applications, Web Services

1. Introduction

With 24 official EU languages and many additional ones, multilingualism, cross-lingual and cross-cultural communication in Europe as well as an inclusive EU Digital Single Market can only be enabled and firmly established through Language Technologies (LTs) (Rehm, 2016). The European LT industry is dominated by hundreds of SMEs and a few large players. Many are world-class, with technologies that outperform the global players. However, European LT business is also fragmented – by nation states, languages, domains and sectors (Vasiljevs et al., 2019), significantly holding back its impact. In addition, many European languages are severely under-resourced and, thus, in danger of digital language exinction (Rehm and Uszkoreit, 2012; Kornai, 2013; Rehm et al., 2014; Rehm et al., 2016a), which is why there is an enormous need for a European LT platform as a unifying umbrella (Rehm and Uszkoreit, 2013; Rehm et al., 2016b; STOA, 2017; Rehm, 2017; Rehm and Hegele, 2018; European Parliament, 2018).[1] The project European Language Grid (ELG; 2019-2021) addresses this fragmentation by establishing the ELG as the primary platform and marketplace for the European LT community, both industry and research. The ELG is a scalable cloud platform, providing access to hundreds of commercial and non-commercial LTs for all European languages, including running tools and services as well as data sets and resources. The ELG will enable the commercial and non-commercial European LT community to upload their technologies and data sets into the ELG in an easy and efficient way, to deploy them, and to connect with other resources.

2. Approach and Methodology

The European LT community has been demanding a dedicated LT platform for years. The ELG project's ambition is to establish the European Language Grid as the primary platform for industry-relevant LT in Europe, bringing together and uniting a network of European experts and concentrating on *commercial* and *non-commercial LTs*, both *functional* and *non-functional* (corpora, lexicons, data sets etc.). A related goal is to establish the ELG as the primary market place for the fragmented European LT landscape to connect demand and supply, strengthening Europe's position in this field. The ELG is meant to enable the whole European LT community to upload their services and data sets, to deploy them and to connect with, and make use of those resources made available by others. *The ELG is meant to be a shared platform for the whole European LT community*, enabling LT provider companies to grow and benefit from scaling up and also companies who want to integrate LT into their products or services.

3. The European Language Grid

In the following, we describe the architecture of the ELG platform (Section 3.1.), including the repository catalogue (Section 3.2.) as well as the user interface (Section 3.3.). Section 3.4. provides more details on the functional services available in the ELG. The language resources are discussed in Section 3.5., followed by a description of the ELG community (Section 3.6.). Section 3.7. provides an overview of the open calls for pilot projects.

3.1. Technical Architecture of the Infrastructure

ELG is a scalable platform with an interactive web user interface and backend components and APIs. It offers access to various language resources as well as functional LT services, i. e., LT tools that have been containerised and wrapped with the ELG LT Service API. ELG's integrated functional services can be used through APIs or through the web interface. The architecture is separated into three layers (Figure 1), i. e., the *base infrastructure* (Kintzel et al., 2019; Moritz et al., 2019), the *platform backend* (Piperidis et al., 2019; Labropoulou et al., 2019) and the *platform frontend* (Melnika et al., 2019a; Melnika et al., 2019b).

The *base infrastructure* is operated on a Kubernetes cluster in the data centre of a European provider located in Berlin, Germany. All infrastructural components run as Docker containers in this cluster.

[1] This article is a shortened version of (Rehm et al., 2020a).

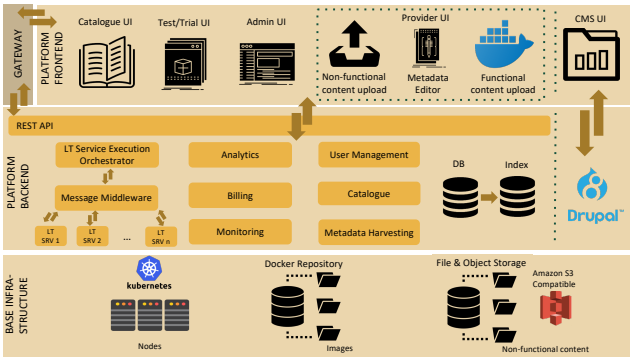

Figure 1: Technical architecture of the ELG

The *platform backend* contains the ELG catalogue, i. e., the metadata records of functional services, non-functional resources but also the entries of organisations (e. g., companies, universities) and other stakeholders, as well as service types, languages and other types of information. Stakeholders will be able to register themselves in this catalogue for increased reach and visibility. Users will be able to filter and search for organisations, services, data sets and more, by languages, service types, domains, and countries. The platform backend layer also includes the LT Service Execution Server/orchestrator that offers a common REST API for calling integrated functional services.

The *platform frontend* layer consists of UIs for the different types of ELG users, e. g., LT providers, potential buyers and ELG system administrators (Section 3.3.). These include catalogue UIs, test/trial UIs for functional services, provider UIs for uploading/registering functional services etc.

One of the key concepts of the architecture is the use of containers to encapsulate all components, settings and libraries of an individual LT service in one self-contained unit. Docker is currently the most widely used technology for containerisation. For individual LT services, Docker images can be built locally by their respective providers and ingested into the ELG, where they can be started, terminated and scaled out on demand.

3.2. Catalogue Structure – Metadata Schema

The ELG catalogue contains all entities of interest to users (Section 3.6.), appropriately indexed and described so that they can easily find and select the resources that meet their requirements and deploy them, as well as visualise the LT domain activities, stakeholders and resources with specific criteria (e. g., service type, language, etc.). All entities are described in compliance with the ELG-SHARE metadata schema (Labropoulou et al., 2019; Labropoulou et al., 2020). The schema builds upon, consolidates and updates previous activities, especially the META-SHARE schema and its profiles (Gavrilidou et al., 2012; Piperidis et al., 2018; Labropoulou et al., 2018), taking into account the ELG user requirements (Melnika et al., 2019a), recent developments in the (meta)data domains (e. g., FAIR, data and software citation recommendations, Open Science movement, etc.), and the need for establishing a common pool of resources through exchange mechanisms with collaborating projects and initiatives (Section 3.6.).

3.3. User Interface

To identify the user scenarios and requirements, we defined the main groups of ELG users: (1) *Content providers* – companies, research organisations or public institutions with tools, services, or data that can be provided through the ELG; (2) *Developers and integrators* – companies and research institutions interested in using LT services, tools, or data in their applications; (3) *Information seekers* – users interested in information about LT, data or events; (4) *Information providers* – organisations or individuals who wish to provide information about events, trainings etc.; (5) *Casual visitors*; (6) *ELG platform administrators*.

Angular and Typescript are used for developing the ELG front-end. The Angular Material components are implemented as an adjustable theme that can be tuned to the designer's specifications. For example, a theme has primary and secondary colours that, once set, will be used throughout all interface elements. The website design is based on the Single Page Application (SPA) principle. To enable the flexible management of content and information within the ELG website, we integrated Drupal. As the ELG front-end is a monolithic SPA, the CMS does not have a dedicated public front-end. Instead, Drupal serves different menus and page contents using REST services and JSON-HAL.

3.4. Functional Services

The European LT market is very broad and varied, with many different providers of many different classes of services and tools, exposed as many different APIs and data formats. One of ELG's primary goals is to attempt to bring more order to this varied landscape by identifying classes of related services and providing generic APIs for each class. From the outset the project has identified a number of broad classes that share much in common:

Machine Translation (MT): services that take text in one language and translate it into text in another language, possibly with additional metadata associated with each segment (sentence, phrase, etc.). This class can include (seemingly unrelated) services such as summarisation, where the summary can be viewed as a "translation" of the original text.

Information Extraction (IE): services that take text and annotate it with metadata on specific segments. This class can cover a wide variety of services from basic NER through to complex sentiment analysis and domain-specific tools.

Automatic Speech Recognition (ASR): services that take audio as input and produce text (e. g., a transcription) as output, possibly with metadata associated with each segment.

Other clusters are emerging as the project considers more services for integration, for example text-to-speech, text classification, alignment, and translation quality estimation. An aspiration for the platform is to provide services of all classes for all official EU languages and for other EU and non-EU languages that are of strategic interest within the EU. The current prototype has so far integrated seven IE and text analysis tools plus a multilingual dependency parser supporting 60 languages, five ASR services (one supporting three languages and another supporting two), 14 MT services (six languages into English, English into eight other languages) and text-to-speech in four languages.

3.5. Data Sets and Language Resources

The ELG consortium has defined an LR identification and sharing strategy. It starts by liaising with and capitalizing on existing activities to ingest LRs into the ELG, which often requires some sort of negotiation with the owners/providers to obtain the rights to do so. We currently focus on providers who are part of the consortium (ELDA/ELRA) and recent well-known activities such as ELRC-SHARE (Lösch et al., 2018; Piperidis et al., 2018) and META-SHARE (Piperidis, 2012; Piperidis et al., 2014). ELG is working both on data integration procedures, where metadata compliance is key for the exchange of data and metadata descriptions (Section 3.2.), and on the implementation of market place-related features, such as upload/download, licensing, billing, payment, etc.

LR modalities covered are text (corpora, lexicons, etc.), speech/audio, video/audiovisual, images/OCR, sign language, and others. We currently work on a classification matrix that includes LR types, modalities and languages with the goal of analysing the status of existing LRs and LTs. About 220 additional repositories have been located so far, which will increase the current numbers as the exploration and ingestion of LRs is progressing. ELG will approach resource users and suppliers to offer them an additional market channel and will look into both research organisations and companies that build or use commercial or non-commercial LRs. As a first step, over 650 LRs from ELRA, ELRC-SHARE and META-SHARE are being prepared for integration into the ELG, to be completed for the first release of the ELG platform (April 2020). The following resources will be comprised.

3.6. Stakeholders and Community

ELG aims to respond to the challenge of Europe's fragmented European LT landscape (Vasiljevs et al., 2019), both with regard to industry and research. We address this issue by bringing together all stakeholders under a common umbrella platform, which is why outreach, communication and further community building play a crucial role in ELG. Our main target users are described in Section 3.6.1. In addition, we have been setting up two community instruments, the National Competence Centres (Section 3.6.2.) and the European LT Council (Section 3.6.3.).

3.6.1. Key Stakeholders

ELG caters, first, for *commercial LT providers* who want to showcase their products, services and their own organisation. We want to provide *the* marketplace for European LT, which requires a broad geographical, technological and sectorial representation of, ideally, all relevant European provider companies. Organisations will be able to claim (or delete) their record through the ELG user interface so that they can take over maintenance and populate their ELG page. *Research centres and universities* are also LT providers but their interest is not a monetary but a research-driven one. This group provides data sets or smaller tools including rudimentary, experimental services that have evolved from research projects, rather than fully-fledged services ready for production and monetisation. *LT users* are the most diverse target group. It includes organisations who want to make use of LT, students doing research for a paper or job seekers. Members of this group can be on the lookout for information, try to find free services or be potential buyers. ELG is collaborating or in the process of setting up collaborations with several relevant projects and initiatives that have similar goals, such as AI4EU, ELRC, BDVA, CLAIRE, CLARIN, HumanE-AI and META-NET (Rehm et al., 2020b). The *participants in the pilot projects*, funded through the ELG open calls, are also key stakeholders. ELG will test the platform and demonstrate its usefulness with the help of 15-20 pilot projects that receive financial support through the project budget. The results of these pilots will be fed back into the ELG platform.

3.6.2. National Competence Centres

ELG set up 32 National Competence Centres (NCCs) to extend the reach of the platform and initiative. They were selected based on their involvement in relevant community initiatives. The fact that all NCCs have good connections to major local industry sectors while being part of academic organisations, guarantees independence from economic interests while ensuring sufficient outreach into commercial fields to serve the purpose of ELG. The NCCs function as bridges between the national and regional markets and the ELG, both as a platform and project. They provide information about stakeholders, services, data sets, resources and technologies from the given region. They know the language(s) and the political as well as economic situation in their countries and are represented in regional networks.

3.6.3. European LT Council

ELG is also initiating a second new body, the European LT Council (LTC), as a pan-European group in which strategic LT-related matters can be discussed and coordinated. While the main purpose of the NCCs is to support the mission of the ELG project, the main goal of the LTC is to support and represent the European LT community. The LTC is meant to be a forum that enables easy and efficient communication and coordination at the European level, specifically with regard to ongoing and emerging international and also national activities relating to LT research, development and innovation. The LTC fosters the coordination and strategic as well as political discussion, representing all relevant stakeholder groups. It will prepare strategic recommendations, especially geared towards national and European administrations and funding agencies.

3.7. Open Calls for Pilot Projects

ELG will provide close to 30% of its overall budget for a set of 15-20 small scale demonstrator pilot projects in the form of grants awarded after a call for proposals. The pilot projects will broaden ELG's portfolio by developing missing services or solutions that support underrepresented languages. At the same time, they will demonstrate the ELG's usefulness as a technology platform, especially with regard to sectors of high commercial or societal impact. The results of the projects will be made available through the ELG. LT tools or services will be integrated into the ELG itself and made generally available under defined licensing conditions. The main objective of the open calls is to support SMEs that have long-term potential to either (a) contribute

services, tools or data sets to the ELG to increase its coverage or (b) develop applications using LTs available in the ELG. Financial support will be awarded to selected applicants following an open, transparent and expert-evaluation based selection process. Each proposal will be evaluated by three independent experts for the following criteria: (a) objective fit; (b) technical approach; (c) business, integration and dissemination plan; (d) budget adequacy; and (e) team. The first call will be published in March 2020, and the second in September 2020.

4. Sustainability through a Legal Entity

Achieving the intended scale of the ELG requires a high availability and performance of the overall system, service level agreements (SLAs) for the services as well as billing and support facilities. These create various costs, that can only be covered adequately through a sustainable, long-term operational model. Costs include cloud hosting and bandwidth, personnel costs for operations, development, accounting, marketing, support and management, legal consulting (SLAs, GDPR, contracts etc.), office space, computers, electricity etc. ELG is meant to be a sustainable activity. To achieve this goal, we need to identify ways to cover the incurred costs on a long-term basis. We will establish a legal entity by approx. Q1/2021. Among the options are a for profit or non-profit company, a professional stakeholder association and a foundation.

There are various potential ingredients of a future ELG business and operations plan. These include online ads (for companies, services, conferences etc.), sponsored content (e. g., first search result, clearly marked as "sponsored"), i. e., sponsored services, data sets, or companies, among others. The ELG legal entity can also offer training events, tutorials or webinars for a fee for commercial players, while keeping them free for academia. ELG conferences may include registration fees for delegates from industry, also offering sponsorship packages for companies. Consulting services around ELG and language-centric AI can be offered. If we decide to establish a professional business association, membership fees could be part of the business plan. Project grants can be used to sustain part of the operation. Additionally, the hosting of commercial LT services, models or data sets can be associated with a certain fee, while the results of publicly funded research can be made available for free, but the hosting costs would need to be covered nonetheless. In that regard, ELG could function as the secondary or maybe even as the primary dissemination channel for research projects or for companies that develop LT. Part of the ELG business model could also include the brokering of commercial LT services for a fee, with a split between the service owner and ELG as the broker. ELG could also function as a paid hoster for whole service or data repositories.

5. Related Work

Research Projects, Platforms, Initiatives All in all, we have collected more than 30 projects, platforms and initiatives that are, in one way or another, relevant for ELG (Rehm et al., 2020a). They share at least one of the following goals with ELG: 1) they provide a collection of LT/NLP tools or data sets; 2) they provide a unified platform, which, underneath, harvests metadata records of data sets or services or tools from distributed sources; 3) they provide a sharing platform for the exchange of tools or data sets among stakeholders.

Global Technology Enterprises Many of the global technology enterprises offer a wide range of different processing services, beyond language, including cloud and compute resources, storage, different types of databases, data analytics, and also more engineering-related services such as encryption, development and deployment. Among these are offerings by Amazon, especially AWS and Comprehend, Microsoft Azure Cognitive Services (Del Sole, 2018), the Google Cloud Platform and the IBM Cloud (Kochut et al., 2011). Google has recently (Sept. 2018) released a dedicated search platform for datasets.

6. Conclusions and Next Steps

It has repeatedly been argued that Europe should by no means outsource its multilingual communication and language challenge to providers from other continents since the European demands are so unique and complex (Rehm and Uszkoreit, 2013; Rehm, 2017; Rehm et al., 2020b). Instead, Europe should make use of its own excellent LT community. One of the obstacles to be overcome along the way is the creation of a shared platform for the whole community. The ELG will foster language technologies *for Europe* built *in Europe*, tailored to our languages and cultures and to our societal and economical demands, benefitting the European citizen, society, innovation and industry. There is currently no other scalable cloud platform that can play the role as a joint marketplace and broker for such a broad variety of services and data sets as we have foreseen for the ELG.

At the time of writing, the three-year ELG project is at the end of its first year, which has already seen the first public demo of a fully functional miminum viable product of the ELG platform at META-FORUM 2019. Work in all three ELG areas is progressing at a fast pace. The next major milestones will include launching the first open call in March 2020 and, at the same time, making available the first version of the ELG platform to interested parties. This version will also include the first batches of functional services and data sets. The second open call will be published in September 2020, coinciding with the second release of the platform, services and data sets. The third and final release of the platform (including additional services and data sets) is foreseen for the last quarter of 2021. In 2020 and 2021 we will organise two more annual ELG conferences that will also include NCC and LTC meetings. At the end of 2021, a new legal entity will take over the further development and maintenance of the ELG platform. With regard to upcoming funding programmes on the European level, we foresee ELG to play a number of roles, especially as the main data and service provision and dissemination platform for the European LT and language-centric AI community (Rehm et al., 2020b) in Horizon Europe and Digital Europe Programme but also in national funding initiatives.

Funding notice

The work presented in this paper has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 825627 (European Language Grid) and from the German Federal Ministry of Education and Research (BMBF) through the project QURATOR (Wachstumskern no. 03WKDA1A).

7. Bibliographical References

Del Sole, A. (2018). Introducing Microsoft Cognitive Services. In *Microsoft Computer Vision APIs Distilled*, pages 1–4. Springer.

European Parliament. (2018). Report on language equality in the digital age. http://www.europarl.europa.eu/doceo/document/A-8-2018-0228_EN.html, September. (2018/2028(INI)). Committee on Culture and Education (CULT), Committee on Industry, Research and Energy (ITRE); Rapporteur: Jill Evans.

Gavrilidou, M., Labropoulou, P., Desipri, E., Piperidis, S., Papageorgiou, H., Monachini, M., Frontini, F., Declerq, T., Francopoulo, G., Arranz, V., and Mapelli, V. (2012). The META-SHARE Metadata Schema for the Description of Language Resources. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC 2012)*. European Language Resources Association (ELRA).

Hummel, J., Wetzel, M., Kranias, L., and Magnusdottir, G. (2016). LTI Cloud Business Model Report, April. MLi Deliverable D4.3. MLi: Towards a MultiLingual Data & Services Infrastructure.

Kintzel, F., Moritz, M., and Rehm, G. (2019). Requirements and Architectural Specification of the Base Infrastructure, April. ELG Deliverable D1.1. ELG: European Language Grid.

Kochut, A., Deng, Y., Head, M. R., Munson, J., Sailer, A., Shaikh, H., Tang, C., Amies, A., Beaton, M., Geiss, D., et al. (2011). Evolution of the IBM Cloud: Enabling an enterprise cloud services ecosystem. *IBM Journal of Research and Development*, 55(6):7–1.

Kornai, A. (2013). Digital Language Death. *PLoS ONE*, 8(10). https://doi.org/10.1371/journal.pone.0077056.

Labropoulou, P., Galanis, D., Lempesis, A., Greenwood, M., Knoth, P., Eckart de Castilho, R., Sachtouris, S., Georgantopoulos, B., Martziou, S., Anastasiou, L., Gkirtzou, K., Manola, N., and Piperidis, S. (2018). OpenMinTeD: A Platform Facilitating Text Mining of Scholarly Content. In *WOSP 2018 Workshop Proceedings, Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, pages 7–12, Miyazaki, Japan. European Language Resources Association (ELRA).

Labropoulou, P., Gkirtzou, K., Deligiannis, M., Galanis, D., Gavriilidou, M., Piperidis, S., Rehm, G., Moritz, M., and Silva, A. G. (2019). Metadata schema, August. ELG Deliverable D2.3. ELG: European Language Grid.

Labropoulou, P., Gkirtzou, K., Gavriilidou, M., Deligiannis, M., Galanis, D., Piperidis, S., Rehm, G., Berger, M., Mapelli, V., Arranz, V., Choukri, K., Backfried, G., Pérez, J. M. G., and Silva, A. G. (2020). Making Metadata Fit for Next Generation Language Technology Platforms: The Metadata Schema of the European Language Grid. Submitted to LREC 2020.

Lösch, A., Mapelli, V., Piperidis, S., Vasiļjevs, A., Smal, L., Declerck, T., Schnur, E., Choukri, K., and Genabith, J. V. (2018). European Language Resource Coordination: Collecting Language Resources for Public Sector Multilingual Information Management. In Nicoletta Calzolari (Conference chair), et al., editors, *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, Miyazaki, Japan, May 7-12, 2018. European Language Resources Association (ELRA).

Melnika, J., Lagzdiņš, A., Siliņš, U., Skadins, R., and Vasiļjevs, A. (2019a). Requirements and Design Guidelines, June. ELG Deliverable D3.1. ELG: European Language Grid.

Melnika, J., Vasiljevs, A., Skadins, R., and Lagzdins, A. (2019b). User Requirements and Functional Specifications, April. ELG Deliverable D2.1. ELG: European Language Grid.

Moritz, M., Kintzel, F., Elsholz, E., Rehm, G., and Roberts, I. (2019). Base Infrastructure (first release), May. ELG Deliverable D1.2. ELG: European Language Grid.

Piperidis, S., Papageorgiou, H., Spurk, C., Rehm, G., Choukri, K., Hamon, O., Calzolari, N., Gratta, R. D., Magnini, B., and Girardi, C. (2014). META-SHARE: One Year After. In Nicoletta Calzolari (Conference Chair), et al., editors, *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, Reykjavik, Iceland, May. European Language Resources Association (ELRA).

Piperidis, S., Labropoulou, P., Deligiannis, M., and Giagkou, M. (2018). Managing Public Sector Data for Multilingual Applications Development. In Nicoletta Calzolari, et al., editors, *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, Miyazaki, Japan, May. European Language Resources Association (ELRA).

Piperidis, S., Galanis, D., Deligiannis, M., Gkirtzou, K., Labropoulou, P., Rehm, G., Kintzel, F., Moritz, M., Roberts, I., and Bontcheva, K. (2019). Specification of the ELG platform architecture, June. ELG Deliverable D2.2. ELG: European Language Grid.

Piperidis, S. (2012). The META-SHARE Language Resources Sharing Infrastructure: Principles, Challenges, Solutions. In Nicoletta Calzolari (Conference Chair), et al., editors, *Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC'12)*, Istanbul, Turkey, May. European Language Resources Association (ELRA).

Rehm, G. and Hegele, S. (2018). Language Technology for Multilingual Europe: An Analysis of a Large-Scale Survey regarding Challenges, Demands, Gaps and Needs. In Nicoletta Calzolari, et al., editors, *Proceedings of the 11th Language Resources and Evaluation Conference (LREC 2018)*, pages 3282–3289, Miyazaki, Japan, 5. European Language Resources Association (ELRA).

Georg Rehm et al., editors. (2012). *META-NET White Pa-*

per Series "Europe's Languages in the Digital Age". Springer, Heidelberg, New York, Dordrecht, London. 31 volumes on 30 European languages. http://www.meta-net.eu/whitepapers.

Georg Rehm et al., editors. (2013). *The META-NET Strategic Research Agenda for Multilingual Europe 2020*. Springer, Heidelberg, New York, Dordrecht, London.

Rehm, G., Uszkoreit, H., Dagan, I., Goetcherian, V., Dogan, M. U., Mermer, C., Váradi, T., Kirchmeier-Andersen, S., Stickel, G., Jones, M. P., Oeter, S., and Gramstad, S. (2014). An Update and Extension of the META-NET Study "Europe's Languages in the Digital Age". In Laurette Pretorius, et al., editors, *Proceedings of the Workshop on Collaboration and Computing for Under-Resourced Languages in the Linked Open Data Era (CCURL 2014)*, pages 30–37, Reykjavik, Iceland, 5.

Rehm, G., Hajic, J., van Genabith, J., and Vasiļjevs, A. (2016a). Fostering the Next Generation of European Language Technology: Recent Developments – Emerging Initiatives – Challenges and Opportunities. In Nicoletta Calzolari, et al., editors, *Proceedings of the 10th Language Resources and Evaluation Conference (LREC 2016)*, pages 1586–1592, Portorož, Slovenia, May. European Language Resources Association (ELRA).

Rehm, G., Uszkoreit, H., Ananiadou, S., Bel, N., Bielevičienė, A., Borin, L., Branco, A., Budin, G., Calzolari, N., Daelemans, W., Garabík, R., Grobelnik, M., García-Mateo, C., van Genabith, J., Hajič, J., Hernáez, I., Judge, J., Koeva, S., Krek, S., Krstev, C., Lindén, K., Magnini, B., Mariani, J., McNaught, J., Melero, M., Monachini, M., Moreno, A., Odjik, J., Ogrodniczuk, M., Pęzik, P., Piperidis, S., Przepiórkowski, A., Rögnvaldsson, E., Rosner, M., Pedersen, B. S., Skadiņa, I., Smedt, K. D., Tadić, M., Thompson, P., Tufiş, D., Váradi, T., Vasiļjevs, A., Vider, K., and Zabarskaite, J. (2016b). The Strategic Impact of META-NET on the Regional, National and International Level. *Language Resources and Evaluation*, 50(2):351–374. 10.1007/s10579-015-9333-4.

Rehm, G., Berger, M., Elsholz, E., Hegele, S., Kintzel, F., Marheinecke, K., Piperidis, S., Deligiannis, M., Galanis, D., Gkirtzou, K., Labropoulou, P., Bontcheva, K., Jones, D., Roberts, I., Hajic, J., Hamrlova, J., Kacena, L., Choukri, K., Arranz, V., Mapelli, V., Vasiljevs, A., Anvari, O., Lagzdins, A., Melnika, J., Backfried, G., Dikici, E., Janosik, M., Prinz, K., Prinz, C., Stampler, S., Thomas-Aniola, D., Perez, J. M. G., Silva, A. G., Berrio, C., Germann, U., Renals, S., and Klejch, O. (2020a). European Language Grid: An Overview. Submitted to LREC 2020. Marseille, France.

Rehm, G., Marheinecke, K., Hegele, S., Piperidis, S., Bontcheva, K., Hajič, J., Choukri, K., Vasiļjevs, A., Backfried, G., Prinz, C., Pérez, J. M. G., Meertens, L., Lukowicz, P., van Genabith, J., Lösch, A., Slusallek, P., Irgens, M., Gatellier, P., Köhler, J., Bars, L. L., Auksoriūtė, A., Bel, N., Branco, A., Budin, G., Daelemans, W., Smedt, K. D., Garabik, R., Gavriilidou, M., Gromann, D., Koeva, S., Krek, S., Krstev, C., Lindén, K., Magnini, B., Odijk, J., Ogrodniczuk, M., Ras, E., Rögnvaldsson, E., Rosner, M., Pedersen, B. S., Skadiņa, I., Tadić, M., Tufiş, D., Váradi, T., Vider, K., Way, A., and Yvon, F. (2020b). The European Language Technology Landscape in 2020: Language-Centric and Human-Centric AI for Cross-Cultural Communication in Multilingual Europe. Submitted to LREC 2020. Marseille, France.

Georg Rehm, editor. (2016). *Language as a Data Type and Key Challenge for Big Data. Strategic Research and Innovation Agenda for the Multilingual Digital Single Market. Enabling the Multilingual Digital Single Market through technologies for translating, analysing, processing and curating natural language content*. CRACKER and Cracking the Language Barrier federation, 7. Version 0.9. 04 July 2016. Prepared by the Cracking the Language Barrier federation, supported by the EU-funded projects CRACKER and LT_Observatory.

Georg Rehm, editor. (2017). *Language Technologies for Multilingual Europe: Towards a Human Language Project. Strategic Research and Innovation Agenda*. CRACKER and Cracking the Language Barrier federation, 12. Version 1.0. Unveiled at META-FORUM 2017 in Brussels, Belgium, on November 13/14, 2017. Prepared by the Cracking the Language Barrier federation, supported by the EU-funded project CRACKER.

STOA. (2017). Language equality in the digital age – Towards a Human Language Project. STOA study (PE 598.621), IP/G/STOA/FWC/2013-001/Lot4/C2, March 2017. Carried out by Iclaves SL (Spain) at the request of the Science and Technology Options Assessment (STOA) Panel, managed by the Scientific Foresight Unit (STOA), within the Directorate-General for Parliamentary Research Services (DG EPRS) of the European Parliament, March. http://www.europarl.europa.eu/stoa/.

Vasiljevs, A., Choukri, K., Meertens, L., and Aguzzi, S. (2019). Final study report on CEF Automated Translation value proposition in the context of the European LT market/ecosystem. DOI 10.2759/142151. A study prepared for the European Commission, DG Communications Networks, Content & Technology by Crosslang, Tilde, ELDA, IDC.