

Cognitive Technologies

Georg Rehm *Editor*

# European Language Grid

A Language Technology Platform  
for Multilingual Europe



**EUROPEAN  
LANGUAGE  
GRID**

**OPEN ACCESS**



**Springer**

# **Cognitive Technologies**

## **Editor-in-Chief**

Daniel Sonntag, German Research Center for AI, DFKI, Saarbrücken, Saarland,  
Germany

Titles in this series now included in the Thomson Reuters Book Citation Index and Scopus!

The Cognitive Technologies (CT) series is committed to the timely publishing of high-quality manuscripts that promote the development of cognitive technologies and systems on the basis of artificial intelligence, image processing and understanding, natural language processing, machine learning and human-computer interaction.

It brings together the latest developments in all areas of this multidisciplinary topic, ranging from theories and algorithms to various important applications. The intended readership includes research students and researchers in computer science, computer engineering, cognitive science, electrical engineering, data science and related fields seeking a convenient way to track the latest findings on the foundations, methodologies and key applications of cognitive technologies.

The series provides a publishing and communication platform for all cognitive technologies topics, including but not limited to these most recent examples:

- Interactive machine learning, interactive deep learning, machine teaching
- Explainability (XAI), transparency, robustness of AI and trustworthy AI
- Knowledge representation, automated reasoning, multiagent systems
- Common sense modelling, context-based interpretation, hybrid cognitive technologies
- Human-centered design, socio-technical systems, human-robot interaction, cognitive robotics
- Learning with small datasets, never-ending learning, metacognition and introspection
- Intelligent decision support systems, prediction systems and warning systems
- Special transfer topics such as CT for computational sustainability, CT in business applications and CT in mobile robotic systems

The series includes monographs, introductory and advanced textbooks, state-of-the-art collections, and handbooks. In addition, it supports publishing in Open Access mode.

Georg Rehm  
Editor

# European Language Grid

A Language Technology Platform  
for Multilingual Europe

 Springer



*Editor*

Georg Rehm 

Deutsches Forschungszentrum  
für Künstliche Intelligenz GmbH (DFKI)  
Berlin, Germany

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



ISSN 1611-2482

ISSN 2197-6635 (electronic)

Cognitive Technologies

ISBN 978-3-031-17257-1

ISBN 978-3-031-17258-8 (eBook)

<https://doi.org/10.1007/978-3-031-17258-8>

© The Editor(s) (if applicable) and The Author(s) 2023. This book is an open access publication.

**Open Access** This book is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this book are included in the book's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the book's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG  
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

# Foreword

I was proud to have the opportunity to present my report on *Language Equality in the Digital Age* to the European Parliament in 2018 and even prouder to see the overwhelming support it received. It was one of my final achievements as a Member of the European Parliament and I am delighted that it contributed to the groundbreaking work being done on the European Language Grid project. Despite it not being a legislative report, the level of cross-party support it received meant its recommendations could not be ignored.

When I first proposed a report on language equality in the digital age to the European Parliament's Culture and Education Committee it provoked a great deal of interest, as it did in the Industry Committee. This was due to the clear language inequality in Europe but also to the huge opportunities it presented for the digital industries. As both committees laid claim to the report, there was some debate before it was resolved that it would be a Culture Committee report with a written opinion from the Industry Committee. The latter's participation strengthened the report and its impact. It widened the scope to emphasise the importance of the role of private companies alongside public bodies and of facilitating cross border trade in the Digital Single Market.

It was clear from the early days that the European Commission was keen to support the report and take the proposals forward. As a spokesperson for the Commissioner stated in a conference I organised in parliament in September 2018, "You are never so wealthy as when you can speak in your own language". The European Language Equality project is currently developing a roadmap to achieve language equality by 2030, which will be presented to the European Institutions later this year.

Minority languages in particular have most to lose but also most to gain from the digital age, given the right support. Cultural and linguistic diversity depends largely on the technological resources available to all languages.

It was a report by the EU Panel for the Future of Science and Technology, STOA, that sparked the idea of a parliamentary report. STOA highlighted the social and economic consequences of language barriers and the widening of the technological gap. As someone who had long campaigned for equal status for the Welsh language, I was inspired by the potential that a major EU project could offer.

Even though the dominance of a few well-resourced languages in the digital world was obvious, the impact of this on other languages had not been adequately explored. When the discussion began, the interest grew. The increase in technology presented new threats and new opportunities. This was an issue which literally affected everyone, and most notably children growing up in this digital world. The role of education is crucial in teaching and understanding language technologies but also in raising awareness of career opportunities in this industry across Europe.

The European Union itself, of course, could play a major role. The institutional framework for the provision of language technology could be improved considerably. I believed that this was such a crucial issue that it deserved the specific allocation of the portfolio to a European Commissioner. This did not materialise in the appointment of a new Commission following the European elections in 2019, but I believe the proposal should be maintained and could be adopted in future.

The strong support given to my report by the European Parliament was an indication of support for the exciting language equality work that is taking place now. The report proposed a dedicated funding programme for research, development and innovation in language technologies with the aim of closing the gap between European languages.

This suggestion was a direct result of seeing the existing research being done in many countries. Identifying the problem went hand in hand with discovering that there were many individuals and organisations already addressing it and working to overcome it. They had the information and expertise but needed far more support and a higher profile. It was clear that the EU could become a trailblazer in research on digital language technology, given the political will.

As a politician, the rights of minority languages like my own, Welsh, were at the heart of my work for justice and equality. For me, language was not merely a means of communication but central to our culture and identity. The EU claims equality in diversity but when it came to language equality it fell far short. So in my role as a Member of the European Parliament I saw an opportunity to help correct this. I could play my role in parliament but to ensure the report was effective in achieving its aim it needed the input of the experts, the practitioners and the pioneers in this field of work to ensure that it was accurate and informed.

I never fail to be inspired by their work and their dedication and I repeat my thanks to all those who contributed to the success of the *Language Equality in the Digital Age* report and to the remarkable European Language Grid project which established the primary platform for “language technologies *for* Europe built *in* Europe”.

# Preface

The origins of this book date back to 2012. Back then, under the umbrella of the EU Network of Excellence META-NET, we prepared the recommendations and priority research themes specified in the first Strategic Research Agenda (SRA) for the European Language Technology (LT) field in a complex, community-driven process.<sup>1</sup> While the European LT community is quite extensive, with hundreds of commercial and academic organisations working on a large and heterogeneous set of technologies, it is also extremely fragmented with many community members operating only in narrow niches and limited regions, on very specific topics and quite often only taking into account one or two regionally confined languages. Through the META-NET SRA process, we have been able to identify the community’s need for a joint technology platform that brings the European LT community together, that fosters collaboration and synergies, that acts as a marketplace and deployment platform, that functions like the “yellow pages” of the European LT community and through which essentially all European resources, corpora, datasets and grammars as well as tools, services and source code can be discovered and actually used, straight from the platform itself. Back in 2013, in the published META-NET SRA, we called this concept the *European Service Platform for Language Technologies*. The SRA document only contained a rather coarse-grained description of this ambitious technology vision, which has been demanded, for a number of different reasons, by an overwhelming majority of the members of the LT community.

Later on, in the three Strategic Research and Innovation Agendas prepared under the umbrella of the EU project CRACKER (Cracking the Language Barrier; 2015-2017), we refined the notion of the European LT Service Platform and we extended the possible use cases and a large number of LT-driven applications, primarily focusing the multilingual digital single market. Further boosted by the scientific breakthroughs produced in the area of Artificial Intelligence, Machine Learning and Deep Learning, early on applied to LT applications such as Machine Translation, not only the topic of Language Technology but also the vision of a joint European Language Technology Platform became more and more relevant. The topic was mentioned in

---

<sup>1</sup> <http://www.meta-net.eu/sra>



a prominent way in the STOA study *Language equality in the digital age*<sup>2</sup>, commissioned by the European Parliament and also in a European Parliament resolution<sup>3</sup> with the same title, adopted by the European Parliament in a landslide vote in 2018 (cf. Jill Evans' foreword).

Roughly at the same time, in late 2017, we started preparing a project proposal for the Horizon 2020 ICT call, topic ICT-29 a), *European Language Grid*, which fortunately reflected the vast interest within the community in such a platform. After various unsuccessful attempts at coming up with a good title for the proposal, we decided to use the title of the actual call because it fit perfectly. Having passed the evaluation with a positive result, the project started in January 2019. We had an enthusiastic kick-off meeting, exciting hackathons and developed the first prototype of the platform in a fast and agile way. It was first presented to the public at META-FORUM 2019, which took place in Brussels in October that year.

Only a few weeks later, the global SARS-CoV-2 pandemic hit. The whole world was affected and so was our project plan. We were unable to have face-to-face project meetings or additional hackathons, we were unable to organise any on-site workshops with our 32 ELG National Competence Centres as part of the “ELG European Roadshow”. All meetings, including our annual META-FORUM conferences in 2020 and 2021, had to go virtual, which was new to us at first and quickly became the new normal. Recently, in early June 2022, we had our last META-FORUM conference under the umbrella of the ELG EU project. META-FORUM 2022 went back, at least partially, to the *old* normal with approx. 100 participants in the conference centre in Brussels and hundreds more participating remotely.

It was nothing but a pleasure to act as Coordinator of the European Language Grid project and to work together with such a strong and dedicated team. Our original plan in this Innovation Action was already quite ambitious yet we managed to exceed our joint expectations in terms of the technology platform and its features, in terms of the services and resources developed, collected and ingested into the platform, in terms of the acceptance and feedback by the community and also in terms of the various collaborations we conducted with other projects. Many of the features envisioned for the *European Service Platform for Language Technologies* in 2013 are in fact now finally available in the *European Language Grid*, which is, by a large margin, the biggest all-purpose Language Technology platform on the planet covering the whole breadth and technology spectrum of the field.

All of the activities and results produced by the nine partners of the ELG consortium during the project's runtime are described in this book in detail. I would like to thank all consortium partners and team members for their extremely hard and dedicated work towards our common goal of developing and establishing the ELG platform, community and marketplace. In addition, I would like to thank the 15 selected pilot projects for their innovative proposals and the more than 200 organisations who applied for funding through one of our pilot projects. Thanks are also due to the projects ELG collaborated with, especially, in 2021/2022, the European

---

<sup>2</sup> [https://www.europarl.europa.eu/stoa/en/document/EPRS\\_STU\(2017\)598621](https://www.europarl.europa.eu/stoa/en/document/EPRS_STU(2017)598621)

<sup>3</sup> [https://www.europarl.europa.eu/doceo/document/TA-8-2018-0332\\_EN.html](https://www.europarl.europa.eu/doceo/document/TA-8-2018-0332_EN.html)

Language Equality project, the results of which will also be documented in the form of a book in the same series, but also others such as Bergamot, COMPRISE, ELITR, EMBEDDIA, Gourmet, Prêt-à-LLOD, AI4EU, HumanE AI Net, VISION, TAILOR, WeVerify, NTEU, Microservices at your Service, MAPA, QURATOR, PANQURA, SPEAKER and many others.

This book is the definitive documentation of the EU project European Language Grid.<sup>4</sup> I would like to thank all colleagues from the ELG consortium and also from the ELG pilot projects wholeheartedly for the chapters they contributed, without which this book would not have been possible.

While this book can only cover the results achieved during the project's runtime (January 2019 until June 2022), the ELG initiative will continue. In the second half of 2022 we will establish a legal entity that will take over maintenance and operation of the platform. We hope that ELG will serve its many purposes and, among others, address the stark community fragmentation and contribute to digital language equality in Europe, functioning indeed as one joint umbrella platform for the whole European LT community. Furthermore, while none of these can be considered a direct follow-up just yet, in a few projects (including OpenGPT-X, NFDI4DataScience and AI as well as the EU projects DataBri-X and SciLake) we will have the opportunity to continue our work with and on the ELG platform.

Berlin, July 2022

*Georg Rehm*

**Acknowledgements** The European Language Grid EU project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 825627.

---

<sup>4</sup> <https://european-language-grid.readthedocs.io> provides more details with regard to technical aspects of the ELG platform. The online documentation is actively maintained and kept up to date.

# Contents

<b>1</b>	<b>European Language Grid: Introduction</b> .....	<b>1</b>
	Georg Rehm	
1	Overview and Context .....	1
2	The European Language Grid EU Project .....	4
3	Beyond the ELG EU Project .....	6
4	Summary of this Book .....	7
4.1	Part I: ELG Cloud Platform .....	7
4.2	Part II: ELG Inventory of Technologies and Resources ..	8
4.3	Part III: ELG Community and Initiative .....	8
4.4	Part IV: ELG Open Calls and Pilot Projects .....	8
	References .....	9

## Part I ELG Cloud Platform

<b>2</b>	<b>The European Language Grid Platform: Basic Concepts</b> .....	<b>13</b>
	Stelios Piperidis, Penny Labropoulou, Dimitris Galanis, Miltos Deligiannis, and Georg Rehm	
1	Introduction .....	13
2	Overview of the ELG Platform .....	14
2.1	Catalogue .....	15
2.2	Repository of Language Resources and Technologies ...	16
2.3	Running Language Technology Cloud Services .....	16
3	User Types and User Model .....	17
4	Architecture .....	18
5	Catalogue Contents and Metadata Model .....	19
6	Publication Life Cycle .....	24
7	ELG and the FAIR Principles .....	25
8	Related Platforms and Infrastructures .....	27
9	Conclusions .....	30
	References .....	30
	Appendix .....	34

<b>3</b>	<b>Using the European Language Grid as a Consumer</b> . . . . .	<b>37</b>
	Ian Roberts, Penny Labropoulou, Dimitris Galanis, Rémi Calizzano, Athanasia Kolovou, Dimitris Gkoumas, Andis Lagzdīņš, and Stelios Piperidis	
1	Introduction . . . . .	37
2	Web-based Interface . . . . .	39
2.1	Viewing the Catalogue . . . . .	39
2.2	Searching the Catalogue . . . . .	40
2.3	Viewing Metadata Records and Resources . . . . .	43
2.4	Consumer's Grid . . . . .	46
2.5	Try out UIs for Language Technology Services . . . . .	48
3	Public REST APIs . . . . .	50
3.1	Accessing and Using the Catalogue . . . . .	51
3.2	Downloading a Resource . . . . .	51
3.3	Language Technology Service Public API . . . . .	51
4	Python SDK for Users . . . . .	55
4.1	Browsing the Catalogue . . . . .	55
4.2	Downloading a Resource . . . . .	56
4.3	Obtaining an Access Token . . . . .	56
4.4	Calling Language Technology Services . . . . .	57
5	User Authentication . . . . .	58
6	Licensing and Billing . . . . .	59
7	Consumer-Related Functionalities in ELG and other Platforms . . . . .	60
7.1	Catalogue and Repository Functionalities . . . . .	60
7.2	Language Technology Service Execution . . . . .	61
8	Conclusions . . . . .	64
	References . . . . .	65
<b>4</b>	<b>Contributing to the European Language Grid as a Provider</b> . . . . .	<b>67</b>
	Dimitris Galanis, Penny Labropoulou, Ian Roberts, Miltos Deligiannis, Leon Voukoutis, Katerina Gkirtzou, Rémi Calizzano, Athanasia Kolovou, Dimitris Gkoumas, and Stelios Piperidis	
1	Introduction . . . . .	67
2	Adding Resources to the ELG Platform . . . . .	68
2.1	Creating Metadata Records . . . . .	69
2.2	Uploading and Managing Data Files . . . . .	71
2.3	Managing Catalogue Entries . . . . .	72
3	Validating and Publishing Metadata Records . . . . .	74
4	Entity-Type Specific Requirements . . . . .	74
4.1	ELG-compatible Services . . . . .	75
4.2	ELG-hosted Resources . . . . .	88
4.3	Metadata Records for External Resources . . . . .	89
5	Provider-Related Functionalities in ELG and other Platforms . . . . .	89
5.1	Metadata Requirements . . . . .	89
5.2	Provider User Interface and Metadata User Interface . . . . .	90

5.3	Try Out User Interface .....	91
5.4	Helper Tools for Packaging Resources .....	91
5.5	Packaging Data Resources .....	91
6	Conclusions .....	92
	References .....	92
<b>5</b>	<b>Cloud Infrastructure of the European Language Grid .....</b>	<b>95</b>
	Florian Kintzel, Rémi Calizzano, and Georg Rehm	
1	Introduction .....	95
2	Cloud Infrastructure .....	96
2.1	Kubernetes and Cloud Native .....	97
2.2	Storage .....	98
2.3	Software Repositories .....	98
2.4	Container Registries .....	99
3	Installation .....	100
3.1	ELG Charts .....	101
3.2	Third-Party Charts .....	102
4	Scalability of LT Tools and Services .....	104
4.1	Implementation .....	104
4.2	Use Cases .....	104
5	Conclusions .....	105
	References .....	106
<b>6</b>	<b>Interoperable Metadata Bridges to the wider Language Technology Ecosystem .....</b>	<b>107</b>
	Penny Labropoulou, Stelios Piperidis, Miltos Deligiannis, Leon Voukoutis, Maria Giagkou, Ondřej Kořarko, Jan Hajič, and Georg Rehm	
1	Introduction .....	107
2	Approach .....	108
3	Establishing Interoperable Connections: Four Use Cases .....	110
3.1	Use Case 1: OAI-PMH (CLARIN, ELRC-SHARE) .....	110
3.2	Use Case 2: Custom API (Hugging Face) .....	112
3.3	Use Case 3: General Catalogues (Zenodo) .....	113
3.4	Use Case 4: Collaborative Initiatives (ELE, ELG) .....	115
3.5	Summary of Use Cases .....	117
4	Implementing Metadata Interoperability .....	117
4.1	ELG Metadata Schema – Relaxed Version .....	118
4.2	Publication Policies for Imported Metadata Records .....	118
5	Interoperability across Repositories .....	119
5.1	Technical Interoperability across Repositories .....	120
5.2	Semantic Interoperability across Repositories .....	120
5.3	Minimal Metadata Requirements .....	121
5.4	Duplicate Resources .....	122
6	Conclusions .....	122
	References .....	123

## Part II ELG Inventory of Technologies and Resources

<b>7</b>	<b>Language Technology Tools and Services</b> . . . . .	131
	Ian Roberts, Andres Garcia Silva, Cristian Berrio Aroca, Jose Manuel Gómez-Pérez, Miroslav Jánošík, Dimitris Galanis, Rémi Calizzano, Andis Lagzdīņš, Milan Straka, and Ulrich Germann	
1	Introduction . . . . .	131
2	Machine Translation . . . . .	135
3	Automatic Speech Recognition . . . . .	136
3.1	Case Study: Speech Tools from HENSOLDT . . . . .	137
4	Text Analytics . . . . .	138
4.1	Case Study: Cogito Discover from Expert.AI . . . . .	139
4.2	Case Study: GATE from University of Sheffield . . . . .	140
4.3	Case Study: Microservices At Your Service . . . . .	141
5	Other Service Types . . . . .	142
5.1	Pilot Project: Text2TCS, University of Vienna . . . . .	144
5.2	Pilot Project: MKS as LLOD, Coreon . . . . .	145
6	Conclusions . . . . .	146
	References . . . . .	147
<b>8</b>	<b>Datasets, Corpora and other Language Resources</b> . . . . .	151
	Victoria Arranz, Khalid Choukri, Valérie Mapelli, Mickaël Rigault, Penny Labropoulou, Miltos Deligiannis, Leon Voukoutis, and Stelios Piperidis	
1	Introduction . . . . .	151
2	Identification of Language Resources and Repositories . . . . .	152
2.1	Identification by the Consortium . . . . .	153
2.2	Identification by the National Competence Centres . . . . .	153
2.3	Collaboratively Filling the Gaps . . . . .	153
3	Integrating Repositories into ELG . . . . .	155
3.1	Priorities in the Ingestion Work . . . . .	155
3.2	Contributing Language Resources . . . . .	156
4	Procedures to Ingest Language Resources . . . . .	157
4.1	Metadata Conversion . . . . .	157
4.2	Metadata Extraction and Completion . . . . .	158
4.3	Metadata Harvesting . . . . .	160
5	Language Resources in the ELG Catalogue . . . . .	162
6	Language Resources and Legal Issues . . . . .	164
7	Language Resources and Data Management . . . . .	166
8	Conclusions . . . . .	166
	References . . . . .	167

<b>9</b>	<b>Language Technology Companies, Research Organisations and Projects</b> .....	171
	Georg Rehm, Katrin Marheinecke, Rémi Calizzano, and Penny Labropoulou	
1	Introduction .....	171
2	The European Language Technology Landscape .....	174
3	Organisations in the European Language Grid .....	177
3.1	Collecting the Members of the European LT Community	178
3.2	Preparation and Integration of Metadata Records .....	178
3.3	Claiming and Enriching Organisation Pages .....	179
3.4	Organisation Pages in the European Language Grid .....	180
4	Projects in the European Language Grid .....	180
5	Conclusions .....	181
	References .....	184

### Part III ELG Community and Initiative

<b>10</b>	<b>European Language Technology Landscape: Communication and Collaborations</b> .....	189
	Georg Rehm, Katrin Marheinecke, and Jens-Peter Kückens	
1	Introduction .....	189
2	Stakeholders of the European Language Grid .....	191
2.1	Language Technology Providers .....	191
2.2	Language Technology Users .....	192
2.3	Additional Horizon 2020 EU Projects .....	194
2.4	Major European Projects and Initiatives .....	194
2.5	National Competence Centres .....	194
2.6	Public at Large .....	195
3	Communication and Outreach Activities .....	195
3.1	Communication Strategy .....	195
3.2	Communication Campaign .....	196
4	Collaborations with other Projects and Initiatives .....	199
5	Conclusions .....	202
	References .....	202
<b>11</b>	<b>ELG National Competence Centres and Events</b> .....	205
	Katrin Marheinecke, Annika Grützner-Zahn, and Georg Rehm	
1	Introduction .....	205
2	National Competence Centres .....	206
2.1	Tasks and Responsibilities .....	206
2.2	Role and Structure .....	208
2.3	Visibility and Promotion .....	208
2.4	Operational Aspects .....	210
3	Conferences and Workshops .....	210
3.1	META-FORUM Conference Series .....	210
3.2	ELG Workshops .....	214

3.3	Additional Conferences .....	215
4	Conclusions .....	216
	References .....	216
<b>12</b>	<b>Innovation and Marketplace:</b>	
	<b>A Vision for the European Language Grid .....</b>	<b>219</b>
	Katja Prinz and Gerhard Backfried	
1	Introduction .....	219
2	Innovation .....	220
2.1	Significance of Innovation .....	220
2.2	Types of Innovation and Innovation Strategies .....	221
2.3	Open Innovation in the ELG Platform and Marketplace ..	222
3	Multi-sided Marketplace Approach .....	226
3.1	Foundations for a Successful Marketplace .....	228
3.2	ELG Ecosystem of Participants .....	230
3.3	Technical and Practical Aspects .....	231
4	Conclusions .....	231
	References .....	232
<b>13</b>	<b>Sustaining the European Language Grid:</b>	
	<b>Towards the ELG Legal Entity .....</b>	<b>233</b>
	Georg Rehm, Katrin Marheinecke, Stefanie Hegele, Stelios Piperidis, Kalina Bontcheva, Jan Hajič, Khalid Choukri, Andrejs Vasiljevs, Gerhard Backfried, Katja Prinz, Jose Manuel Gómez-Pérez, and Ulrich Germann	
1	Introduction .....	234
2	Long-term Vision and Mission of ELG .....	234
2.1	Mission of the European Language Grid .....	235
2.2	Added Value for Stakeholders .....	236
3	Main Pillars of the Business and Operational Model .....	236
3.1	Expectations by the ELG Consortium's SME Partners ..	237
3.2	Key Aspects of the ELG Legal Entity .....	239
3.3	Assessment of Operational Costs .....	240
3.4	Business Model Canvas .....	241
3.5	Product Portfolio and Revenue Streams .....	245
3.6	Legal Entity Type .....	250
4	Summary and Next Steps .....	252
	References .....	253
<b>Part IV ELG Open Calls and Pilot Projects</b>		
<b>14</b>	<b>Open Calls and Pilot Projects .....</b>	<b>257</b>
	Lukáš Kačena, Jana Hamrlová, and Jan Hajič	
1	Introduction .....	257
2	Organisation of the Open Calls .....	258
2.1	Management Structure and Organisation .....	258



2.2	Timeline	259
2.3	Communication with Stakeholders	259
2.4	Submission Process	260
2.5	Evaluation Process	261
3	Results	262
3.1	Open Call 1	262
3.2	Open Call 2	265
4	Pilot Project Execution	268
5	Conclusions	269
	References	270
<b>15</b>	<b>Basque-speaking Smart Speaker based on Mycroft AI</b>	<b>271</b>
	Igor Leturia, Ander Corral, Xabier Sarasola, Beñat Jimenez, Silvia Portela, Arkaitz Anza, and Jaione Martinez	
1	Overview and Objectives of the Pilot Project	271
2	Mycroft Localisation	272
3	Privacy, Gender and Proximity	272
4	Developments in Basque Speech Technology	273
4.1	ASR Robustness in Noisy Environments	273
4.2	ASR Closed Grammar-based Recognition	274
4.3	Neural Network-based Basque TTS	274
4.4	Gender-neutral Voice	275
5	Conclusions and Results of the Pilot Project	275
	References	276
<b>16</b>	<b>CEFR Labelling and Assessment Services</b>	<b>277</b>
	Mark Breuker	
1	Overview and Objectives of the Pilot Project	277
2	Methodology	278
3	Implementation	279
4	Evaluation	281
5	Conclusions and Results of the Pilot Project	281
	References	282
<b>17</b>	<b>European Clinical Case Corpus</b>	<b>283</b>
	Bernardo Magnini, Begoña Altuna, Alberto Lavelli, Anne-Lyse Minard, Manuela Speranza, and Roberto Zanoli	
1	Overview and Objectives of the Pilot Project	283
2	Corpus Collection and Annotation	284
3	Implementation	285
4	Evaluation	286
5	Conclusions and Results of the Pilot Project	287
	References	288

<b>18</b>	<b>Extracting Terminological Concept Systems from Natural Language Text</b> .....	289
	Dagmar Gromann, Lennart Wachowiak, Christian Lang, and Barbara Heinisch	
1	Overview and Objectives .....	289
2	Methodology .....	290
2.1	Preprocessing .....	290
2.2	Term Extraction .....	291
2.3	Relation Extraction .....	292
2.4	Postprocessing .....	292
3	Evaluation .....	293
4	Conclusions and Results of the Pilot Project .....	293
	References .....	293
<b>19</b>	<b>Italian EVALITA Benchmark Linguistic Resources, NLP Services and Tools</b> .....	295
	Viviana Patti, Valerio Basile, Andrea Bolioli, Alessio Bosca, Cristina Bosco, Michael Fell, and Rossella Varvara	
1	Overview and Objectives of the Pilot Project .....	295
2	Methodology .....	296
2.1	Surveying the EVALITA Tasks .....	296
2.2	The EVALITA Knowledge Graph .....	297
2.3	Anonymisation of Resources .....	298
2.4	Release of Data and Models through ELG .....	299
3	Conclusions and Results of the Pilot Project .....	299
	References .....	299
<b>20</b>	<b>Lingsoft Solutions as Distributable Containers</b> .....	301
	Sebastian Andersson and Michael Stormbom	
1	Overview and Objectives of the Pilot Project .....	301
2	Methodology .....	302
3	Implementation .....	302
4	Evaluation .....	304
5	Conclusions and Results of the Pilot Project .....	305
	References .....	306
<b>21</b>	<b>Motion Capture 3D Sign Language Resources</b> .....	307
	Zdeněk Krňoul, Pavel Jedlička, Miloš Železný, and Luděk Müller	
1	Overview and Objectives of the Pilot Project .....	307
2	Methodology and Experiment .....	308
2.1	Recording Setup .....	309
2.2	Data Annotation .....	309
2.3	Data Post-processing .....	310
2.4	Dataset Parameters .....	311
3	Conclusions and Results of the Pilot Project .....	311
	References .....	312

<b>22</b>	<b>Multilingual Image Corpus</b> .....	313
	Svetla Koeva	
1	Overview and Objectives of the Pilot Project .....	313
2	Methodology .....	314
	2.1 Ontology of Visual Objects .....	314
	2.2 Collection of Images and Metadata .....	315
3	Criteria for the Selection of Images .....	315
	3.1 Generation and Evaluation of Suggestions .....	316
	3.2 Annotation Protocol .....	316
4	Multilingual Classes .....	317
5	Conclusions and Results of the Pilot Project .....	317
	References .....	318
<b>23</b>	<b>Multilingual Knowledge Systems as Linguistic Linked Open Data</b> ...	319
	Alena Vasilevich and Michael Wetzel	
1	Overview and Objectives of the Pilot Project .....	319
2	Making Coreon Data Structure LLOD-compatible .....	320
3	Real-Time Data Access via a SPARQL Endpoint .....	323
4	Conclusions and Results of the Pilot Project .....	323
	References .....	324
<b>24</b>	<b>Open Translation Models, Tools and Services</b> .....	325
	Jörg Tiedemann, Mikko Aulamo, Sam Hardwick, and Tommi Nieminen	
1	Overview and Objectives of the Pilot Project .....	325
2	Increasing Language Coverage .....	326
3	Conclusions and Results of the Pilot Project .....	329
	References .....	329
<b>25</b>	<b>Sign Language Explanations for Terms in a Text</b> .....	331
	Helmut Ludwar and Julia Schuster	
1	Overview and Objectives of the Pilot Project .....	331
2	Methodology .....	332
3	Implementation .....	334
4	Evaluation .....	335
5	Conclusions and Results of the Pilot Project .....	335
	References .....	336
<b>26</b>	<b>Streaming Language Processing in Manufacturing</b> .....	337
	Patrick Wiener and Steffen Thoma	
1	Overview and Objectives of the Pilot Project .....	337
2	Graphical, Flow-based Modeling with Apache StreamPipes .....	338
3	Architecture .....	339
4	Implementation .....	340
5	Conclusions and Results of the Pilot Project .....	341
	References .....	342

<b>27</b>	<b>Textual Paraphrase Dataset for Deep Language Modelling</b> . . . . .	343
	Jenna Kanerva, Filip Ginter, Li-Hsin Chang, Valtteri Skantsi, Jemina Kilpeläinen, Hanna-Mari Kupari, Aurora Piirto, Jenna Saarni, Maija Sevón, and Otto Tarkka	
1	Overview and Objectives of the Pilot Project . . . . .	343
2	Methodology . . . . .	344
3	Implementation . . . . .	345
4	Evaluation . . . . .	345
5	Conclusions and Results of the Pilot Project . . . . .	347
	References . . . . .	348
<b>28</b>	<b>Universal Semantic Annotator</b> . . . . .	349
	Roberto Navigli, Riccardo Orlando, Cesare Campagnano, and Simone Conia	
1	Overview and Objectives of the Pilot Project . . . . .	349
2	Methodology . . . . .	350
3	Implementation . . . . .	350
4	Evaluation . . . . .	352
5	Conclusions and Results of the Pilot Project . . . . .	353
	References . . . . .	353
<b>29</b>	<b>Virtual Personal Assistant Prototype YouTwinDi</b> . . . . .	355
	Franz Weber and Gregor Jarisch	
1	Overview and Objectives of the Pilot Project . . . . .	355
2	Methodology . . . . .	356
	2.1 Use Case 1: Automated Translation of local News . . . . .	357
	2.2 Use Case 2: Secure Communication between Virtual Assistants . . . . .	357
3	Implementation . . . . .	357
4	Conclusions and Results of the Pilot Project . . . . .	359
	References . . . . .	360