EUROPEAN LANGUAGE EQUALITY

# D2.11

# Report from NEM

| | |
|---|---|
| Author | Halid Hrasnica |
| Dissemination level | Public |
| Date | 28-02-2022 |

# About this document

| | |
|---|---|
| Project | European Language Equality (ELE) |
| Grant agreement no. | LC-01641480 – 101018166 ELE |
| Coordinator | Prof. Dr. Andy Way (DCU) |
| Co-coordinator | Prof. Dr. Georg Rehm (DFKI) |
| Start date, duration | 01-01-2021, 18 months |
| Deliverable number | D2.11 |
| Deliverable title | Report from NEM |
| Type | Report |
| Number of pages | 38 |
| Status and version | Final |
| Dissemination level | Public |
| Date of delivery | Contractual: 28-02-2022 – Actual: 28-02-2022 |
| Work package | WP2: European Language Equality – The Future Situation in 2030 |
| Task | Task 2.2 The perspective of European LT users and consumers |
| Author | Halid Hrasnica |
| Reviewers | Natalia Resende, Jiřina Popelíková |
| EC project officers | Susan Fraser, Miklos Druskoczi |
| Contact | European Language Equality (ELE)<br>ADAPT Centre, Dublin City University<br>Glasnevin, Dublin 9, Ireland |
| | Prof. Dr. Andy Way – andy.way@adaptcentre.ie |
| | European Language Equality (ELE)<br>DFKI GmbH<br>Alt-Moabit 91c, 10559 Berlin, Germany |
| | Prof. Dr. Georg Rehm – georg.rehm@dfki.de |
| | http://www.european-language-equality.eu |
| |  |

# Consortium

| | | | |
|---|---|---|---|
| 1 | Dublin City University (Coordinator) | DCU | IE |
| 2 | Deutsches Forschungszentrum für Künstliche Intelligenz GmbH (Co-coordinator) | DFKI | DE |
| 3 | Univerzita Karlova (Charles University) | CUNI | CZ |
| 4 | Athina-Erevnitiko Kentro Kainotomias Stis Technologies Tis Pliroforias, Ton Epikoinonion Kai Tis Gnosis | ILSP | GR |
| 5 | Universidad Del Pais Vasco/ Euskal Herriko Unibertsitatea (University of the Basque Country) | UPV/EHU | ES |
| 6 | CROSSLANG NV | CRSLNG | BE |
| 7 | European Federation of National Institutes for Language | EFNIL | LU |
| 8 | Réseau européen pour l'égalité des langues (European Language Equality Network) | ELEN | FR |
| 9 | European Civil Society Platform for Multilingualism | ECSPM | DK |
| 10 | CLARIN ERIC – Common Language Resources and Technology Infrastructure as a European Research Infrastructure Consortium | CLARIN | NL |
| 11 | Universiteit Leiden (University of Leiden) | ULEI | NL |
| 12 | Eurescom (European Institute for Research and Strategic Studies in Telecommunications GmbH) | ERSCM | DE |
| 13 | Stichting LIBER (Association of European Research Libraries) | LIBER | NL |
| 14 | Wikimedia Deutschland (Gesellschaft zur Förderung freien Wissens e. V.) | WMD | DE |
| 15 | Tilde SIA | TILDE | LV |
| 16 | Evaluations and Language Resources Distribution Agency | ELDA | FR |
| 17 | Expert System Iberia SL | EXPSYS | ES |
| 18 | HENSOLDT Analytics GmbH | HENS | AT |
| 19 | Xcelerator Machine Translations Ltd. (KantanMT) | KNTN | IE |
| 20 | PANGEANIC-B. I. Europa SLU | PAN | ES |
| 21 | Semantic Web Company GmbH | SWC | AT |
| 22 | SIRMA AI EAD (Ontotext) | ONTO | BG |
| 23 | SAP SE | SAP | DE |
| 24 | Universität Wien (University of Vienna) | UVIE | AT |
| 25 | Universiteit Antwerpen (University of Antwerp) | UANTW | BE |
| 26 | Institute for Bulgarian Language "Prof. Lyubomir Andreychin" | IBL | BG |
| 27 | Sveučilište u Zagrebu Filozofski fakultet (Univ. of Zagreb, Faculty of Hum. and Social Sciences) | FFZG | HR |
| 28 | Københavns Universitet (University of Copenhagen) | UCPH | DK |
| 29 | Tartu Ulikool (University of Tartu) | UTART | EE |
| 30 | Helsingin Yliopisto (University of Helsinki) | UHEL | FI |
| 31 | Centre National de la Recherche Scientifique | CNRS | FR |
| 32 | Nyelvtudományi Kutatóközpont (Research Institute for Linguistics) | NYTK | HU |
| 33 | Stofnun Árna Magnússonar í íslenskum fræðum SAM (Árni Magnússon Inst. for Icelandic Studies) | SAM | IS |
| 34 | Fondazione Bruno Kessler | FBK | IT |
| 35 | Latvijas Universitātes Matemātikas un Informātikas institūts (Institute of Mathematics and Computer Science, University of Latvia) | IMCS | LV |
| 36 | Lietuvių Kalbos Institutas (Institute of the Lithuanian Language) | LKI | LT |
| 37 | Luxembourg Institute of Science and Technology | LIST | LU |
| 38 | Università ta Malta (University of Malta) | UM | MT |
| 39 | Stichting Instituut voor de Nederlandse Taal (Dutch Language Institute) | INT | NL |
| 40 | Språkrådet (Language Council of Norway) | LCNOR | NO |
| 41 | Instytut Podstaw Informatyki Polskiej Akademii Nauk (Polish Academy of Sciences) | IPIPAN | PL |
| 42 | Universidade de Lisboa, Faculdade de Ciências (University of Lisbon, Faculty of Science) | FCULisbon | PT |
| 43 | Institutul de Cercetări Pentru Inteligență Artificială (Romanian Academy) | ICIA | RO |
| 44 | University of Cyprus, French and European Studies | UCY | CY |
| 45 | Jazykovedný ústav Ľudovíta Štúra Slovenskej akadémie vied (Slovak Academy of Sciences) | JULS | SK |
| 46 | Institut Jožef Stefan (Jozef Stefan Institute) | JSI | SI |
| 47 | Centro Nacional de Supercomputación (Barcelona Supercomputing Center) | BSC | ES |
| 48 | Kungliga Tekniska högskolan (Royal Institute of Technology) | KTH | SE |
| 49 | Universität Zürich (University of Zurich) | UZH | CH |
| 50 | University of Sheffield | USFD | UK |
| 51 | Universidad de Vigo (University of Vigo) | UVIGO | ES |
| 52 | Bangor University | BNGR | UK |

# Contents

# List of Figures



# List of Tables

# List of Acronyms

| | |
|---|---|
| ELE | European Language Equality *(this project)* |
| ELE Programme | European Language Equality Programme *(the long-term, large-scale funding programme specified by the ELE project)* |
| EU | European Union |
| LT | Language Technology/Technologies |
| NEM | New European Media Initiative |
| NLP | Natural Language Processing |
| SME | Small and Medium-sized Enterprise |
| SRIA | Strategic Research and Innovation Agenda |
| USA | United States of America |

## Abstract

The document presents answers following consultation with members of the NEM Initiative who are considered to be Language Technology users and consumers. 26 responses were received from nine European countries with an additional respondent coming from USA. The answers have been gathered mainly from Information and Communication Technologies, Education, Research, and Media sectors. It can be concluded that proofing tools, translation tools, and search engines are widely used, followed by language learning tools. Slightly lower usage can be observed in speech recognition tools and particularly in parsing tools. We can also conclude that the performance of Proofing, Translation, Search, and Language learning tools is estimated to be good or even excellent, whereas Speech recognition, Parsing, Text summarizing, and Text mining tools are estimated to poor and although in rare cases good performance. The most frequently used tools (used every day) are various search engines, followed by Proofing tools (used frequently or every day), and Translation tools (frequently used), whereas Parsing, Sentiment and opinion analysis, Text summarizing, and Text mining tools are never used. As expected, all the tools are mainly used for or in English. To increase usage of the language tools, the respondents emphasised the need for higher-quality tools for the languages they work with (54%) on the first instance, followed by a wider range of language tools available for people to work with (35%).


# 1. Introduction

This document reports on the findings of a consultation with representatives from the LT users and consumers community, conducted by the EU project European Language Equality (ELE). These results will serve as input for a strategic research, innovation and deployment agenda (SRIA) and roadmap, in order to tackle the striking imbalance between European languages in terms of the support they receive through Language Technologies (LTs) by 2030.

The ELE project sought to collect the views of European LT users and consumers and to consolidate their perspective on the differences in terms of technologies for the languages they work with and of the measures that need to be put in place so that all European languages are equally supported through technology by 2030.

Due to the interdisciplinary nature of the field of Language Technology, which stands at the intersection of Linguistics, Computational Linguistics, Computer Science and Artificial Intelligence, the ELE project brings together diverse groups of stakeholders including researchers, representatives of communities of LT users and consumers, language professionals (e. g., translators, lecturers and professors in the field of Linguistics and Computational Linguistics) and stakeholders from different economic sectors (e. g., banking, health).

Although the methodology and instruments utilised have been common to all ELE consortium members, this report covers and analyses the subset of responses of stakeholders contacted by the New European Media initiative (NEM).

## About NEM

The NEM Initiative (New European Media, former Networked and Electronic Media) is a leading European Network for Media and Creative Industries with the mission to foster the impact of interactive technologies on the future of new media through interaction between Media, Content, Creative industries, Social Media, Broadcasting and Telecom sectors as well as Consumer electronics – all together represented by more than 1,000 members of the NEM Community. More information is available on the NEM website.[1] The NEM Initiative's goal

[1] https://nem-initiative.org

is to develop a common innovation environment for the new European media landscape, where equal access to the application of the newest technologies in respect to media is prioritised. In addition to providing accessibility solutions for various disadvantaged groups, such as those with disabilities, and vulnerable populations such as migrants, this includes technologies for automatic translation and transcription, ensuring that media in Europe can reach everyone independently for languages spoken in particular regions, ensuring broad and equal access to information.

# 2. Methodology and Instruments

## 2.1. Online Survey

The survey addressed to LT users and consumers sought to elicit the respondents' views in a way that facilitates the analysis, consolidation and integration of the collected feedback into the ELE SRIA and roadmap. It had 63 questions in total. Some of the questions depend on previous answers. As a result, a respondent was presented with 30 (minimum) to 63 (maximum) questions, including the "if other" questions. 46 questions were mandatory from which 33 were closed questions (single or multiple choice). Table 1 shows an overview of the types of questions.

| Question types | Mandatory | Optional | Totals |
|---|---|---|---|
| Closed | 20 | 13 | 33 |
| Open-ended | 26 | 4 | 30 |
| Totals | 46 | 17 | 63 |

Table 1: Type of survey questions

The survey was structured in four main parts. If any of the provided answers were not applicable, the respondents had the option to enter a different answer through the option "if other, please specify".

- **Part A. Respondents' profiling:** the first part of the survey included 13 questions for the demographic profiling of respondents with emphasis on characteristics relevant to the task at hand, i. e.,
  - Country respondents are based in
  - Name of the organisation/representative body respondents work for
  - Communities they represent (if applicable)
  - Type of organisation respondents work for
  - Sectors or domains that respondents are active in (if applicable)
  - Role of respondents in the organisation (if applicable)
  - Organisation's estimated revenue (if applicable)
- **Part B. Language coverage:** looked into the European languages the respondents work with and the languages they intend to include in their workflow, i. e.,
  - Languages the organisations, associations, communities, professionals of LT users work with
  - Languages planned to be supported in the short- or medium-term

- **Part C. Evaluation of current situation:** assessed the current situation by asking respondents to evaluate the level of technology support for the official European languages they work with and any minority, regional or lesser used language, i. e.,
  - Differences in availability of LTs between the official European languages they work with and, if applicable, differences in availability of LTs between the minority, regional or lesser-used languages they work with;
  - Gaps perceived in the technologies, tools or applications respondents work with especially in relation to specific languages;
  - Respondents' opinion in relation to performance of LTs with regard to specific languages
- **Part D. Predictions and visions for the future:** respondents are requested to share their needs and wishes for the future of language technologies, i. e.,
  - Policies or instruments that could contribute to speed up the effective deployment of LT in Europe equally for all languages
  - Prediction of future opportunities for LT in basic and applied research (scientific vision) and in innovation and the industry
  - Expectations of the community with regard to the challenges an ELE Programme can address by 2030

**Follow-up:** The last three questions requested the respondent's permission to be contacted for an interview and, given an affirmative answer, their contact details. Respondents were also requested to click on a confirmation question stating "By clicking on 'Submit', I agree that my personal data (email address and/or name) can be used according to the Privacy Policy of the European Language Equality (ELE) project".

The survey was designed, set up and published on the EU Survey platform.[2] The full survey, as published online, is presented in Appendix A (p. 13 ff.).

The survey was distributed by NEM through emails to more than 1,000 community members. It has additionally been advertised through the NEM Initiative websites, LinkedIn page and Twitter account.

The survey was opened on 21 June 2021 and closed on 18 October 2021. In total, 246 responses have been collected, out of which 26 respondents were contacted by NEM. This subset of responses, representing the views of the stakeholders contacted by NEM is analysed in this report.

## 2.2. Interviews

All NEM members were approached to participate in the online ELE survey. In order to ensure a higher number of answers, the NEM Steering Board members were directly contacted by providing the survey questions in a Word format, to further facilitate responses. Only one individual volunteered to be interviewed, and this interview was conducted following the survey structure as well. The interview was conducted with an industry organization.

[2] https://ec.europa.eu/eusurvey/runner/LTusers-consumers

# 3. Analysis of Responses

## 3.1. Survey Responses

### 3.1.1. Respondents' profiling

Most of the responses were received from organisations based in Spain (7, 27%), followed by organisations from France (5, 19%), Germany (4, 15%), Belgium (3, 11%), and Portugal (2, 7.5%). One response came from Croatia, Greece, Italy, and Norway respectively. As previously referenced, one survey was completed from USA. The breakdown of answers is shown in Figure 1.

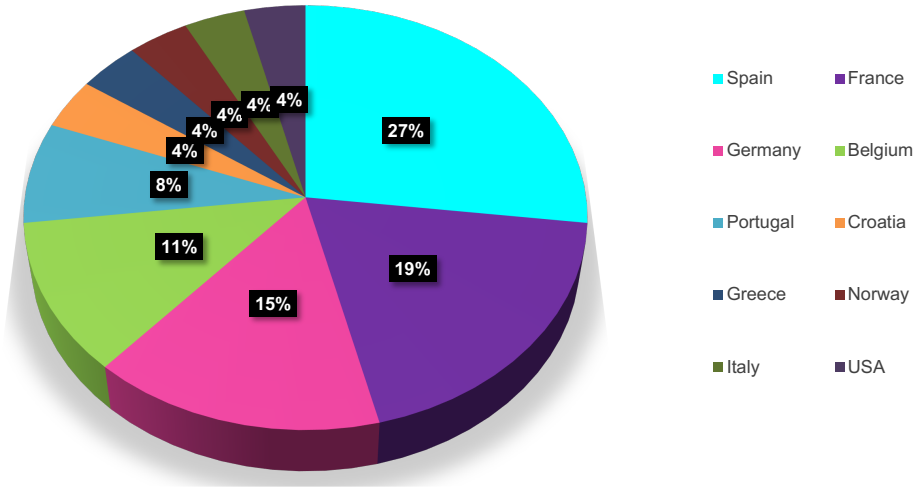



Figure 1: In which country are you based?

The sectors predominantly represented by the respondents from the NEM Initiative are from Information and Communication Technologies and Education (both sectors selected by 10 organisations, Research (9 answers), and Media (8), followed by Digital Humanities, arts, culture and other services (6) and Publishing (4). Two responses were received from each of the Health, Broadcasting, Industry and manufacturing, Business services, and Social Sciences sectors, whereas from Construction, Tourism, accommodation and food services, Transportation, logistics and storage, Energy/green economy /environment, and Public administration received 1 response from each of them. In addition, the responding organisations from NEM indicated Social innovation and Language services as the sectors they belong to (not listed in the initial survey list).

Most of the responses received came from Education/research organisations (12, 46%), followed by Large enterprises (6, 23%) and SMEs (3, 12%). From professional associations and Innovation clusters (latest not listed in the survey) we received 2 answers from each (corresponding to 8%) as well as 1 answer from Independent contractor/ consultant (4%) On size of the responding organisations / number of employees we did not receive enough responses. The breakdown of answers is shown in Figure 2. A more detailed summary of all answers with the breakdown can be found in Appendix B, Table 2 and Table 3.

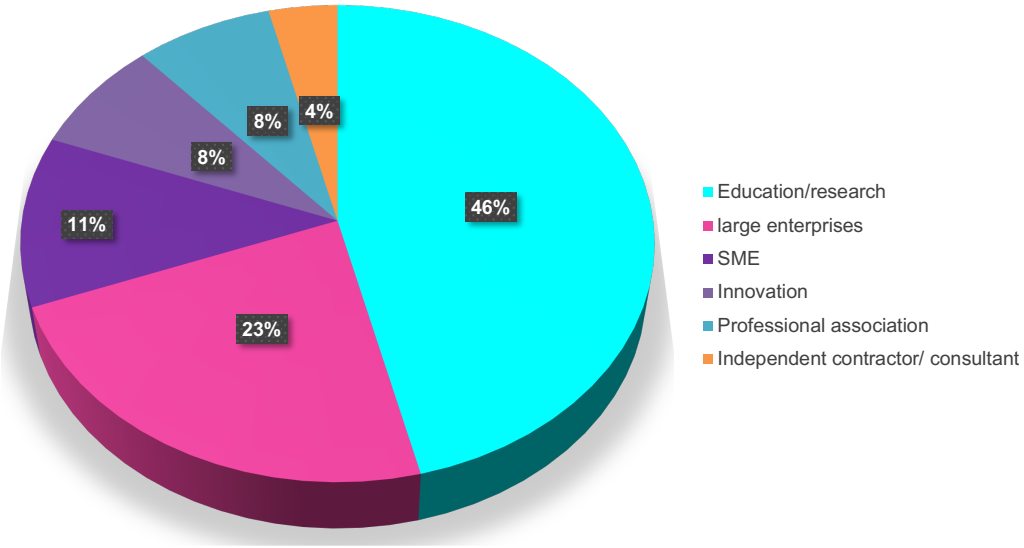



Figure 2: Which of the following best describes the type of organisation you work for?

Around 30% of the respondents were Managing directors/presidents of the organisations and heads of research/innovation or public research units. 23% of respondents were university professors and teachers and the same number of responses was received from Technical/innovation managers. Further answers have been gathered from researchers, Ccommunication consultants, Ttraining translators, and others.

### 3.1.2. Language Coverage

As expected, English is used as one of the working languages in most of the organisations (25 out of 26). Other frequently used languages are Spanish (in 13 organisations), French (11), German (10), and Italian (7), followed by Portuguese (5). Dutch and Greek are used in 3 organisations, Croatian in 2, and Czech, Norwegian, and Slovenian in one organisation.

Catalan is used in two and Basque in one organisation. Thus, percentage of respondents working with minority,regional or lesser-used languages is 3 out of 26 (12%).

Furthermore, Japanese and Chinese (the variant of which is unknown) are used in one organisation. The international corporation are of course using local languages in the countries world-wide they are present in. Figure 3 shows the breakdown of languages selected and Table 4 shows the complete statistics.

English is being included in one of the organisations, French and Spanish in two organisations, and Catalan, Dutch, Slovenian, and Polish in further organisations. Also, the responding organisations mentioned to include further languages in the future, depending on origin of personnel.

### 3.1.3. Evaluation of the Current Situation

While considering answers from the 26 organisations, it can be concluded that proofing tools, translation tools, and searching engines are widely used, followed by language learning tools. Slightly lower usage can be observed in the case of speech recognition tools and in particular

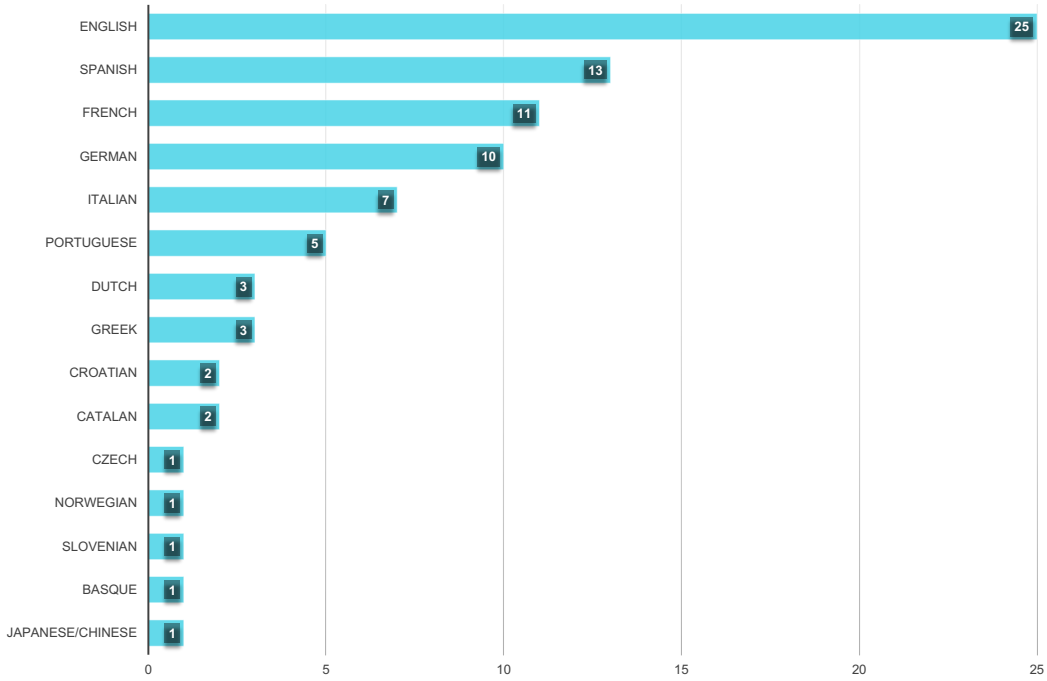



Figure 3: Which of the official European language(s) listed below do you or your organisation work with? if "Other", please specify.

in parsing tools. Related details are provided below. Finally, ***two organisations indicated usage of the Deepl tool, which was not listed within the survey questions***.

- **Proofing tools** are used by all responding organisations
  - Spell checkers (18 )
  - Grammar checkers (18)
  - Autocorrect tools (19)
- Among the **Translation tools**, the most used are Generic translation tools freely available on the web (e. g. Google Translate) in 19 cases, followed by Computer-assisted translation tools (e. g. translation memories) in 8 cases, Terminology management applications (3), and Custom-built translation engines (2).
- Among the **Speech recognition tools**, Voice user interfaces (e. g. Siri, native Android, native iOS, smart speakers [Google home, Alexa, ...], Bose Headphones, Adobe Acrobat reader, Amazon Polly, Chromevox, Wordreference) are used in five cases and Text-to-speech systems (i. e., systems that turn text into speech for reading texts out loud (e. g. Amazon Polly, Adobe Acrobat reader) in six cases. Orange is developing its own tool (Djingo) together with Deutsche Telekom for home services in order to avoid predominance of non-European solutions in the area.
- While considering the **Parsing tools**, dependency or constituency parsing systems to automatically analyse the syntax of written or spoken data (e. g. Stanford NLP's CoreNLP java framework, Stanford NLP Stanza, AllenNLP parsing, UDPipe, MaChAmp) are used in two cases. None of the responding organisations is using the Part-of-speech taggers of any type (e. g. NLTK python library, NLPdotnet).

- As expected, the **Search tools** are used in all responding organisations, with emphasis on the generic search systems freely on the web (e. g. Google search), mentioned in 19 cases, and multilingual search engines (e. g. Wikipedia), mentioned in 18 cases. Web-based question-answering systems (e. g. Stack exchange, StackOverflow, Quora, Google search) are also frequently used, as mentioned in 9 answers. Other tools are used rarely such as cross-language search engines (e. g. eBay, Aliexpress) in four cases, domain-specific search engines (focusing on domain-specific topics, e. g. PubMed, Copernic, CC search) and customer-build search engines (e. g. organisations or vendors create search engines themselves) - both in three cases. Ontology tools for extracting the corresponding domain's terms and the relationships between the concepts that these terms represent in a text (e. g. Robot tool) and private search engines (e. g. Search Encrypt and OneSearch, which use different encryption methods to keep your query private) were both mentioned in two cases, and multimedia search engines (e. g. plantnet, or applications like 'Snooth') in one case. Language-focused search engines (e. g. Baidu) are not used by the responding organisations. Among the customer-build search engines, one of the respondents recommends usage of the Qwant search tool, in order to ensure higher European independence in the area.
- Among the **Language learning tools**, the most used ones are web-based translation search engines (e. g. Linguee) in 11 cases, followed by web-based thesaurus tools reported in six cases and computer-assisted language learning tools (e. g. Duolingo, FluentU, SKELL) in four cases. Intelligent systems to aid and assess reading comprehension (e. g. Whooo's Reading, Storia) are not used by the responding organisations.

Figure 4 shows the categories of tools selected by respondents. The complete statistics of all tools used by respondents is presented in Table 5, Appendix B.

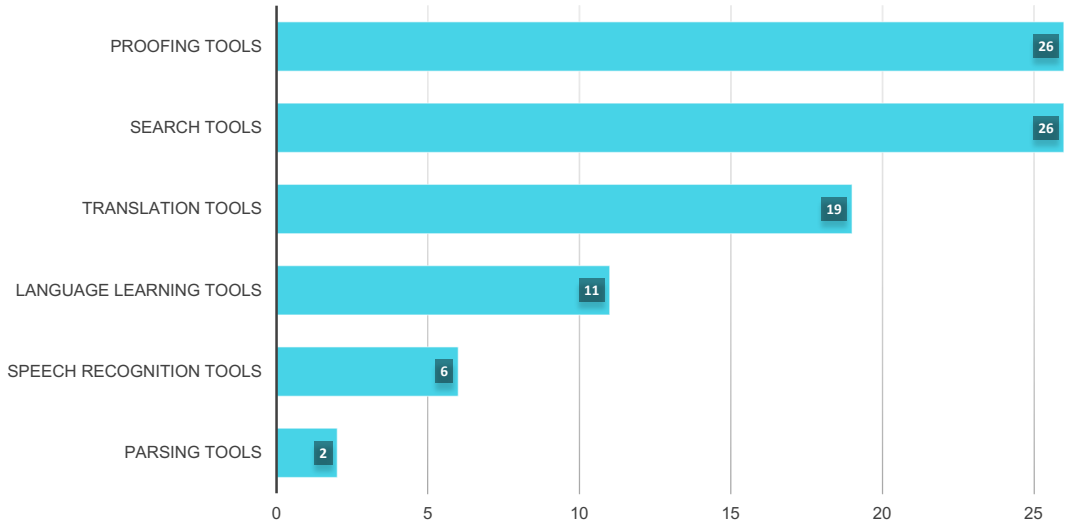



Figure 4: Which language technology tools/applications listed below do you or your organisation use with the official European language(s) you or your organisation work with?

It is interesting to note that 2/3 of the responding organisations did not identify significant gaps in the language technologies that are currently available [3]. The remaining organisations

[3] Please, note that the respondents did not necessarily provide their opinion about all listed tools and answer the questions for the tools they have some knowledge about or experience using

identified the following gaps:

- Amount and variety of available language technology related applications is missing or not available in high quality for Basque, French and Polish, and to a certain extent German, Greek, Italian and Spanish - in principle all except English.
- Adaptability of the language technology tools to various operation systems should be improved, in particular for minority languages.
- An interesting observation was received on Sign Language and the need of its inclusion in the language technologies discussion, directly contributing to improvement of overall accessibility of information systems at large.

We can conclude that the performance of Proofing, Translation, Search, and Language learning tools is estimated to be good or even excellent, whereas Speech recognition, Parsing, Text summarizing, and Text mining tools are estimated to be poor and in some cases good performance. It is interesting to mention that none of the tools is estimated to perform poorly. We have to note, however, that the number of answers for Parsing, Text summarizing, and Text mining tools is much lower, indicating that the respondents do not have much experience working with these tools.

Figure 5 presents answers between 1 and 4 received on technology support for languages the respondents are working with (where 1 = *very poor*, 2= *poor*, 3= *good*, 4= *excellent*). As expected, the best level of support was reported for English, with most of the respondents claiming to receive excellent support, whereas for French, Spanish, German, and Portuguese the technology support is estimated to be good. For Spanish, however, some opinions are rather negative, indicating poor or no support. For Italian, the opinions range between poor and excellent support. For the two minority languages (Catalan and Basque) the answers indicate excellent or good support. Figure 5 shows the mean scores (1-4) for the level of LT support per EU official language.

Based on the one answer received for each of the following languages, we can conclude that good support is available for Dutch and Norwegian, poor support for Croatian and Polish, very poor support for Greek and Romanian.

Regarding frequency of the language tools usage, we note that the most frequently used tools (used every day) are various search engines, followed by Proofing tools (used frequently or every day), and Translation tools (frequently used).

We can observe that Speech recognition and Language learning tools are rarely or never used, whereas Parsing, Sentiment and opinion analysis, Text summarizing, and Text mining tools are never used.

As expected, all the tools are mainly used in English. Lower, but still significant, usage can be observed for French, German, and Spanish. A non-negligible level of usage is reported for Catalan, Dutch, Italian, and Portuguese. For other languages, LTs are used only rarely; it should be noted, however, that this is always dependent on the languages used in the individual organisation.

We have also received some responses which mention usage of the tools for non-European languages, in particular for translation.

### 3.1.4. Predictions and Visions for the Future

Among the responses on the need to use resources in order to increase usage of the language tools, the respondents emphasised the need for higher-quality tools for the languages they work with (14) on the first instance, followed by a wider range of language tools for the languages people work with (9) and More training of personnel dealing with such tools (6).

The answers and ideas gathered on the question about the tools or applications which are currently not available for the working languages are summarised below:

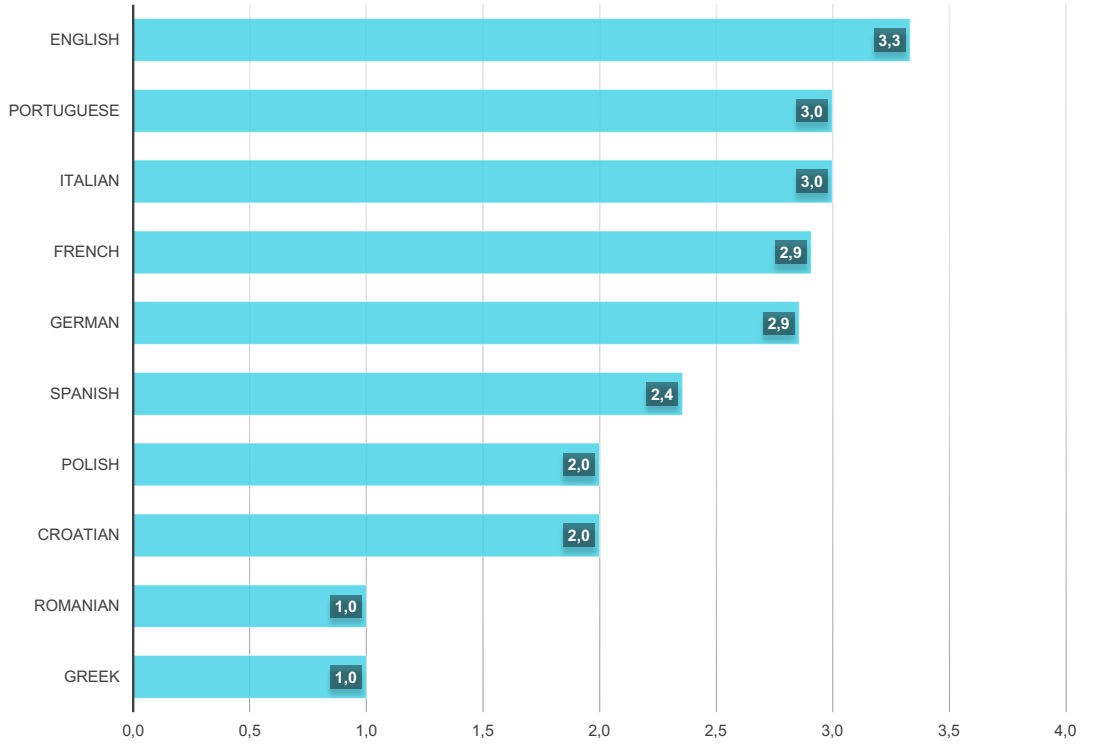



Figure 5: Please choose the option that best describes the level of language technology support for the official European language(s) you or your organisation work with.

- Visio-conference real time translation
- Writing suggestion tools
- Voice recognition transferring automatically to writing
- Automatic translation (text-to-text, speech-to-text, text-to-speech, especially but not exclusively for quality subtitling on videos). The ultimate tool is the “universal translator” known from star trek. Some devices are already available on the market, such as https://vasco-electronics.com/translators/vasco-translator-m3.html
- More language training pairs
- Better sentiment analysis tools (even as simple as word sentiment lists!)
- Speech recognition for Catalan, better grammar checking for Catalan
- Microsoft translator in Basque

Most of the respondents agreed that in the next 10 years, there will be higher quality language tools for all the languages that concern them, including minority languages (11), 8 of them even strongly agreed to this statement, whereas 2 respondents disagreed and others could not say/decide (5).

Most of the respondents even strongly agree that in the next 10 years, there will be a wider range of language tools for European Languages (13), 5 of them agree, whereas 2 respondents disagree. Others cannot say/decide.

On the question of whether language technology tools will help prevent the loss of linguistic diversity in the next 10 years, 8 agreed and 3 even strongly agreed, whereas 4 respondents (15%) disagreed. Others could not say/decide. Figure 6 shows the breakdown of answers.

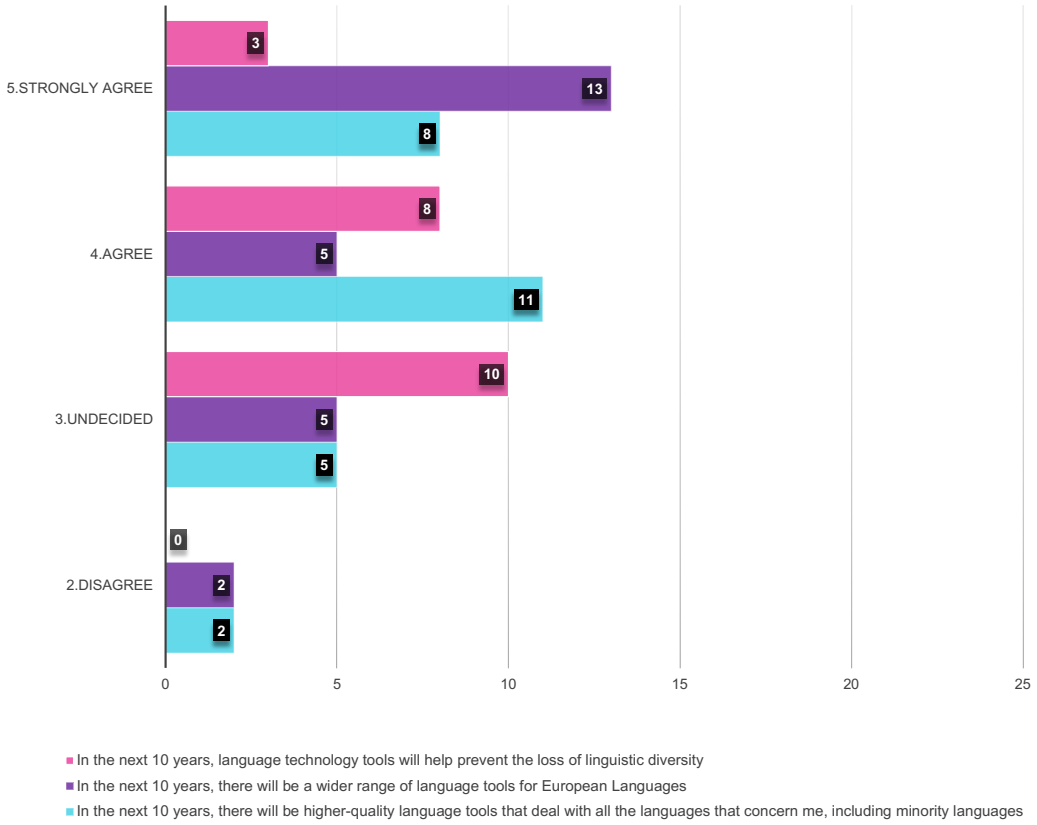



Figure 6: Please indicate the best option that describes your vision for the future of languages technology

The most relevant benefits of improving technologies for the languages people work with are summarised below:

- Increase individuals' exposure to these languages - 10.38%
- Prevent minority/regional languages from disappearing - 12.46%
- Increase the number of speakers of those languages, including minority/regional languages - 3.12%
- Improve communication between native speakers - 12.46%
- Improve literacy for minority/regional languages - 9.35%
- Enhance the communication capabilities of people with disabilities - 13.50%
- Increase engagement with social, leisure and work activities in their own languages - 7.27%
- Improve online trade in countries where those languages are spoken - 13.50%

- Improve offline trade (i. e., not e-commerce) in countries where those languages are spoken - 8.31%

Further suggestions received from the respondents call for support for communication among people speaking different languages and improvement of second language knowledge.

### 3.2. Analysis of Interviews

One interview conducted, as mentioned above, was carried out along the survey questions with an industry organization.

## 4. Conclusions

The document presents responses following consultation with members of the NEM Initiative who are deemed to be Language Technology users and consumers. 26 reponses were received, most of them from Spain and France, followed by Germany, Belgium, and Portugal as well as other countries. Among the responding organisations, besides English, which is the most used working language as expected, Spanish, French, German, Italian, and Portuguese are considered to be frequently used. Catalan and Basque are the only two minority languages which were referenced.

**Language tools usage** When considering responses from the 26 organisations, it can be concluded that the proofing tools, translation tools, and searching engines are widely used, followed by the language learning tools. Slightly lower usage can be observed for speech recognition tools with parsing tools being particularly low. Two organisations indicated usage of the Deepl tool, which was not listed within the survey questions, as well as other tools in development, aimed at reducing the predominance of non-European solutions in the domain.

**Technology gaps** Interestingly, 2/3 of the responding organisations did not identify significant gaps in the currently available language technologies. The remaining organisations identified gaps in the amount and variety of available language technology related applications, which is missing or not available with the certain quality for different languages, however not for English, adaptability of the language technology tools to various operation systems, in particular for minority languages. An interesting observation was received on Sign language and the need of its inclusion in the language technologies discussion.

**Performance of the existing tools** We can conclude that the performance of Proofing, Translation, Search, and Language learning tools is estimated to be good or even excellent, whereas Speech recognition, Parsing, Text summarizing, and Text mining tools are estimated to poor and in rare cases good performance. However, number of answers received for Parsing, Text summarizing, and Text mining tools is rather lower, indicating that the respondents do not have that much experience in working with.

**Language support by the tools** Furthermore, in most of the cases excellent support by language technologies was indicated for English, whereas for French, Spanish, German, and Portuguese the technology support is estimated to be good. For mother languages, the technology support seems to be very low. It is interesting to mention that for the two minority languages (Catalan and Basque) the answers indicate excellent or good support. Es expected, all the tolls are mainly used for or in English, where lower but still significant usage of the tools can be observed for French, German, and Spanish. Not negligible usage of the tolls is reported for Catalan, Dutch, Italian, and Portuguese.

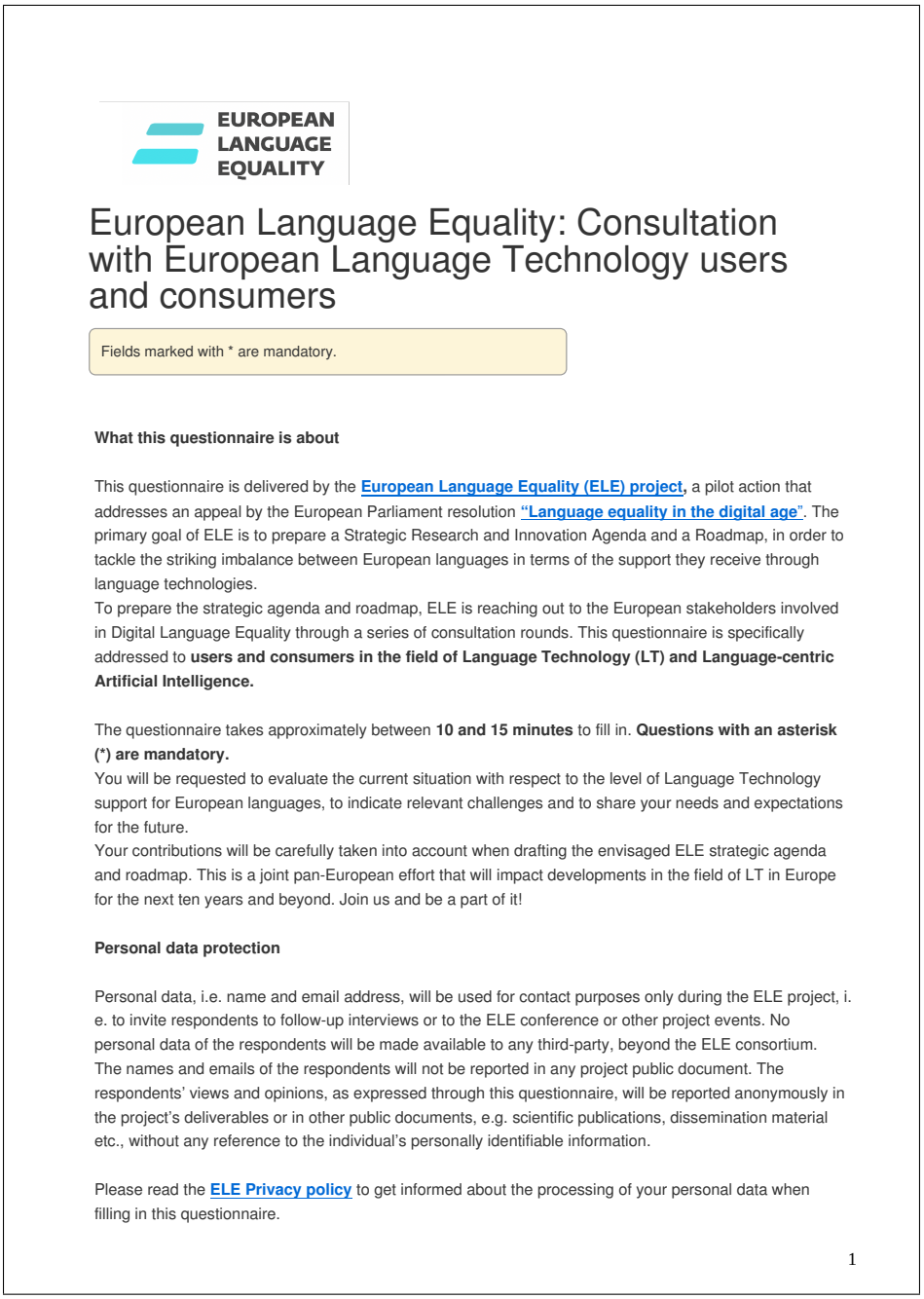
**How to increase the tools usage** To increase usage of the language tools, the respondents emphasised the need for Higher-quality tools for the languages they work with (54%) on the

first instance, followed by a wider range of language tools available for people to work with (35%), and More training of personnel dealing with such tools (23%).

**Outlook** Most of the respondents agree that in the next 10 years, there will be higher quality language tools that deal with all the languages that concern them, including minority languages (42%), 31% even strongly agree to this statement, whereas 8% respondents disagree and others cannot say/decide (19%). Most of the respondents even strongly agree or agree that in the next 10 years, there will be a wider range of language tools for European Languages (69%), whereas 8% disagree. On the question if in the next 10 years, language technology tools will help prevent the loss of linguistic diversity, 43% would agree or strongly agree, whereas 15% or respondents disagree.

# A. LT users and consumers survey

Figures 7 to 24 show the complete LT research and developers survey.



EUROPEAN LANGUAGE EQUALITY

## European Language Equality: Consultation with European Language Technology users and consumers

Fields marked with * are mandatory.

**What this questionnaire is about**

This questionnaire is delivered by the **European Language Equality (ELE) project,** a pilot action that addresses an appeal by the European Parliament resolution **"Language equality in the digital age"**. The primary goal of ELE is to prepare a Strategic Research and Innovation Agenda and a Roadmap, in order to tackle the striking imbalance between European languages in terms of the support they receive through language technologies.
To prepare the strategic agenda and roadmap, ELE is reaching out to the European stakeholders involved in Digital Language Equality through a series of consultation rounds. This questionnaire is specifically addressed to **users and consumers in the field of Language Technology (LT) and Language-centric Artificial Intelligence.**

The questionnaire takes approximately between **10 and 15 minutes** to fill in. **Questions with an asterisk (*) are mandatory.**
You will be requested to evaluate the current situation with respect to the level of Language Technology support for European languages, to indicate relevant challenges and to share your needs and expectations for the future.
Your contributions will be carefully taken into account when drafting the envisaged ELE strategic agenda and roadmap. This is a joint pan-European effort that will impact developments in the field of LT in Europe for the next ten years and beyond. Join us and be a part of it!

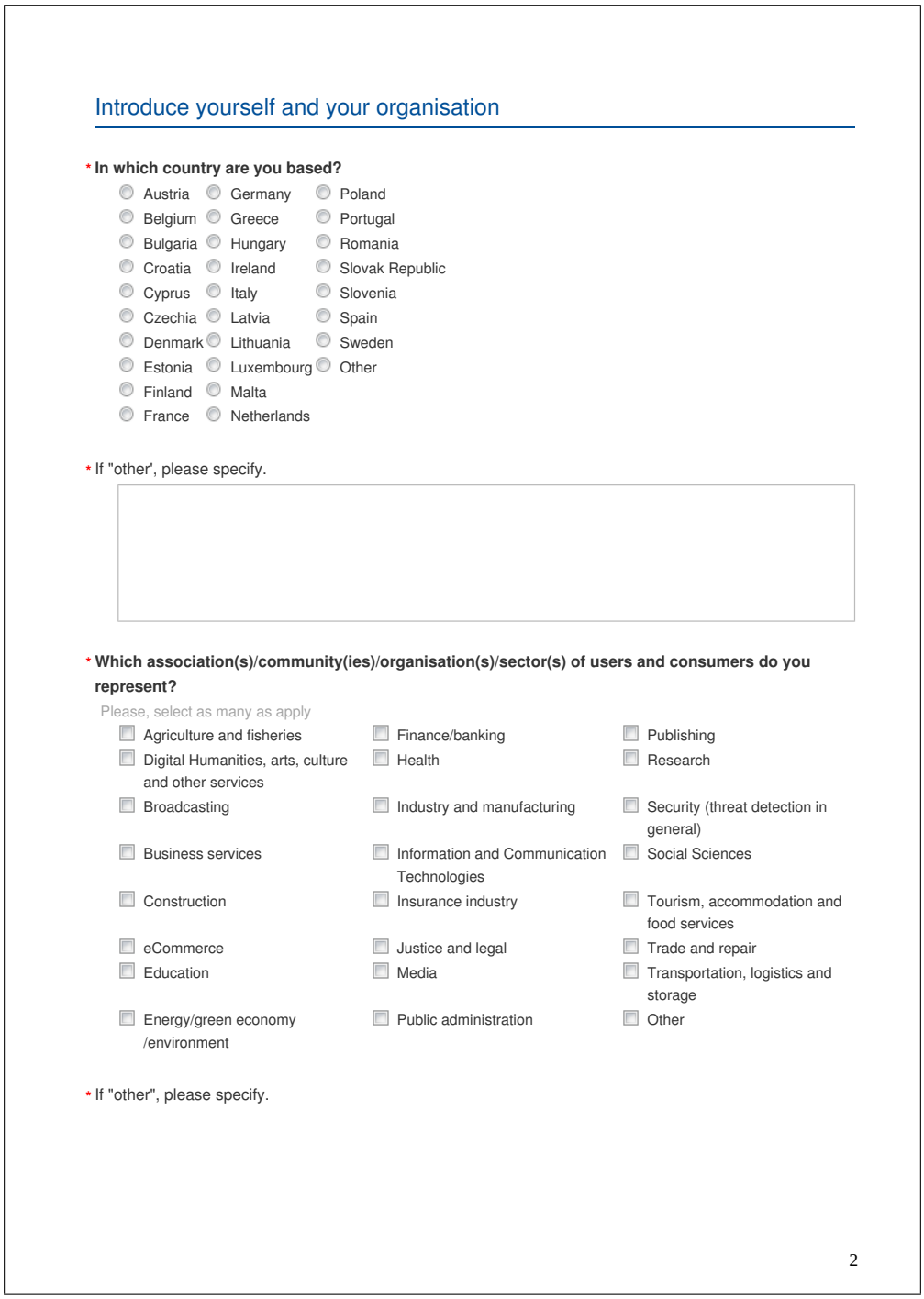
**Personal data protection**

Personal data, i.e. name and email address, will be used for contact purposes only during the ELE project, i. e. to invite respondents to follow-up interviews or to the ELE conference or other project events. No personal data of the respondents will be made available to any third-party, beyond the ELE consortium. The names and emails of the respondents will not be reported in any project public document. The respondents' views and opinions, as expressed through this questionnaire, will be reported anonymously in the project's deliverables or in other public documents, e.g. scientific publications, dissemination material etc., without any reference to the individual's personally identifiable information.

Please read the **ELE Privacy policy** to get informed about the processing of your personal data when filling in this questionnaire.

1

Figure 7: Full survey as published (page 1/18)



## Introduce yourself and your organisation

* **In which country are you based?**

- Austria
- Belgium
- Bulgaria
- Croatia
- Cyprus
- Czechia
- Denmark
- Estonia
- Finland
- France
- Germany
- Greece
- Hungary
- Ireland
- Italy
- Latvia
- Lithuania
- Luxembourg
- Malta
- Netherlands
- Poland
- Portugal
- Romania
- Slovak Republic
- Slovenia
- Spain
- Sweden
- Other

* If "other', please specify.

* **Which association(s)/community(ies)/organisation(s)/sector(s) of users and consumers do you represent?**

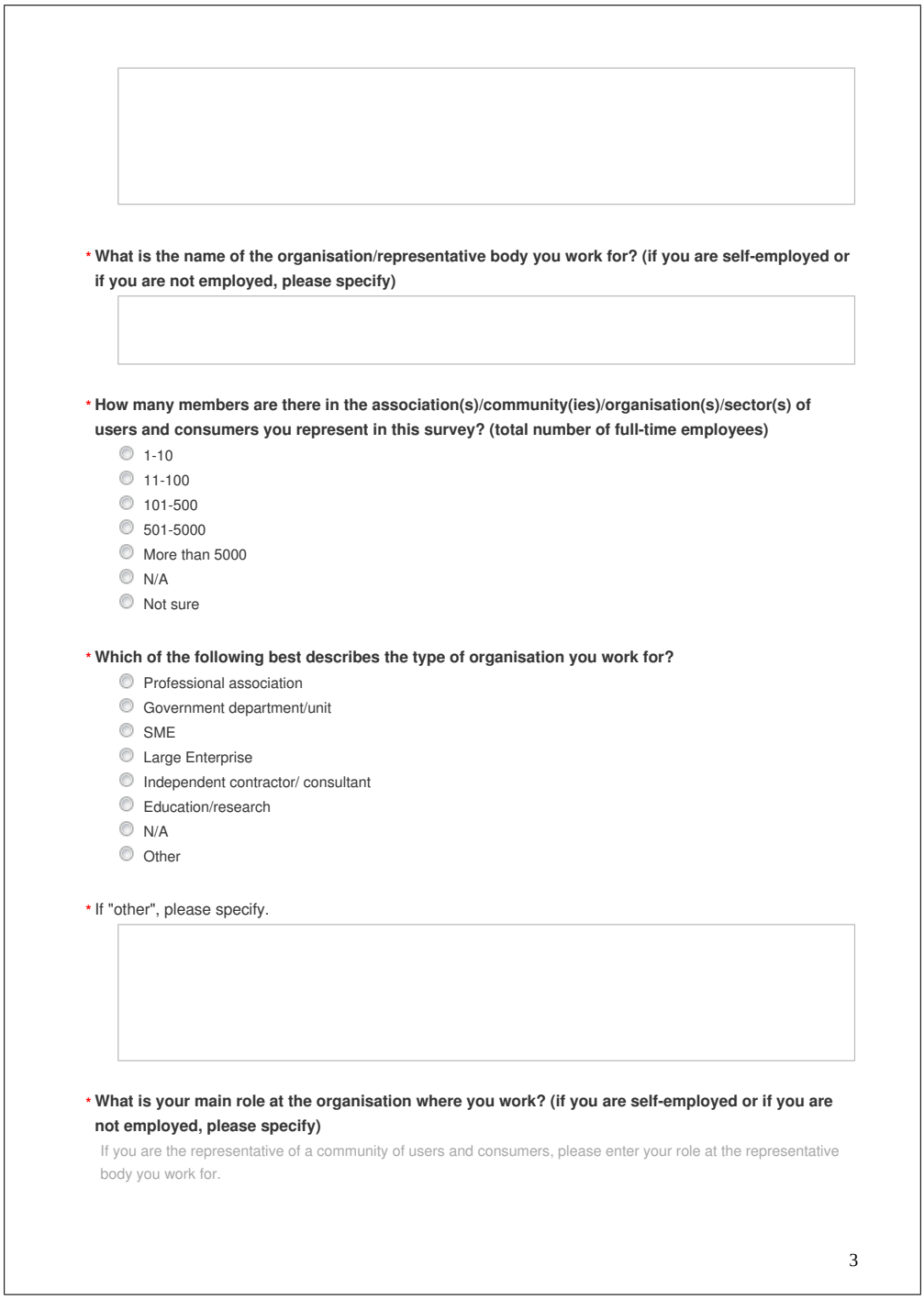
Please, select as many as apply

- Agriculture and fisheries
- Broadcasting
- Business services
- Construction
- eCommerce
- Education
- Energy/green economy /environment
- Finance/banking
- Health
- Industry and manufacturing
- Information and Communication Technologies
- Insurance industry
- Justice and legal
- Media
- Public administration
- Publishing
- Research
- Security (threat detection in general)
- Social Sciences
- Tourism, accommodation and food services
- Trade and repair
- Transportation, logistics and storage
- Other

* If "other", please specify.

2

Figure 8: Full survey as published (page 2/18)



* What is the name of the organisation/representative body you work for? (if you are self-employed or if you are not employed, please specify)

* How many members are there in the association(s)/community(ies)/organisation(s)/sector(s) of users and consumers you represent in this survey? (total number of full-time employees)

- 1-10
- 11-100
- 101-500
- 501-5000
- More than 5000
- N/A
- Not sure

* Which of the following best describes the type of organisation you work for?

- Professional association
- Government department/unit
- SME
- Large Enterprise
- Independent contractor/ consultant
- Education/research
- N/A
- Other

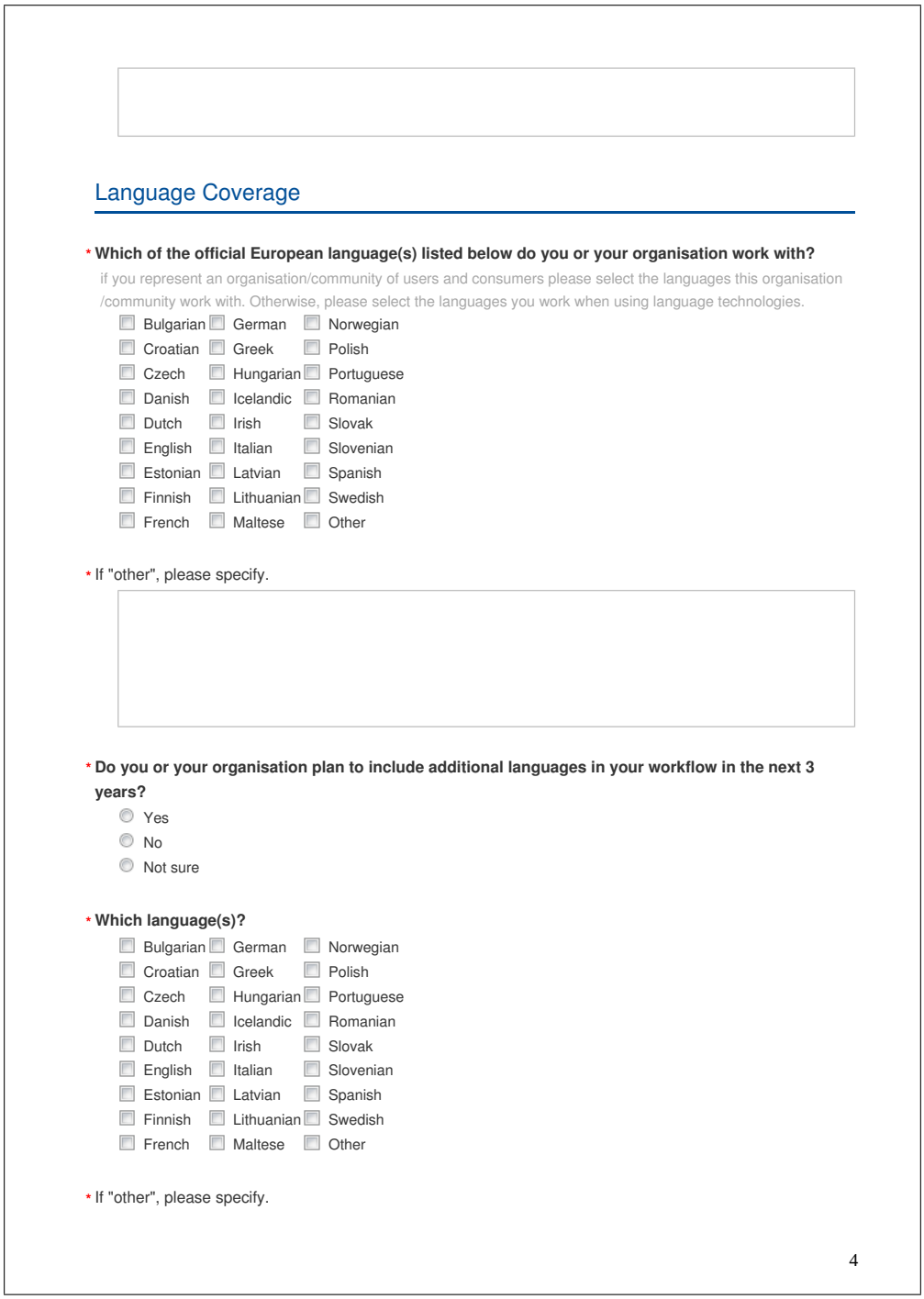
* If "other", please specify.

* What is your main role at the organisation where you work? (if you are self-employed or if you are not employed, please specify)

If you are the representative of a community of users and consumers, please enter your role at the representative body you work for.

3

Figure 9: Full survey as published (page 3/18)



## Language Coverage

* **Which of the official European language(s) listed below do you or your organisation work with?**

if you represent an organisation/community of users and consumers please select the languages this organisation /community work with. Otherwise, please select the languages you work when using language technologies.

- ☐ Bulgarian
- ☐ Croatian
- ☐ Czech
- ☐ Danish
- ☐ Dutch
- ☐ English
- ☐ Estonian
- ☐ Finnish
- ☐ French
- ☐ German
- ☐ Greek
- ☐ Hungarian
- ☐ Icelandic
- ☐ Irish
- ☐ Italian
- ☐ Latvian
- ☐ Lithuanian
- ☐ Maltese
- ☐ Norwegian
- ☐ Polish
- ☐ Portuguese
- ☐ Romanian
- ☐ Slovak
- ☐ Slovenian
- ☐ Spanish
- ☐ Swedish
- ☐ Other

* If "other", please specify.

* **Do you or your organisation plan to include additional languages in your workflow in the next 3 years?**

- ○ Yes
- ○ No
- ○ Not sure

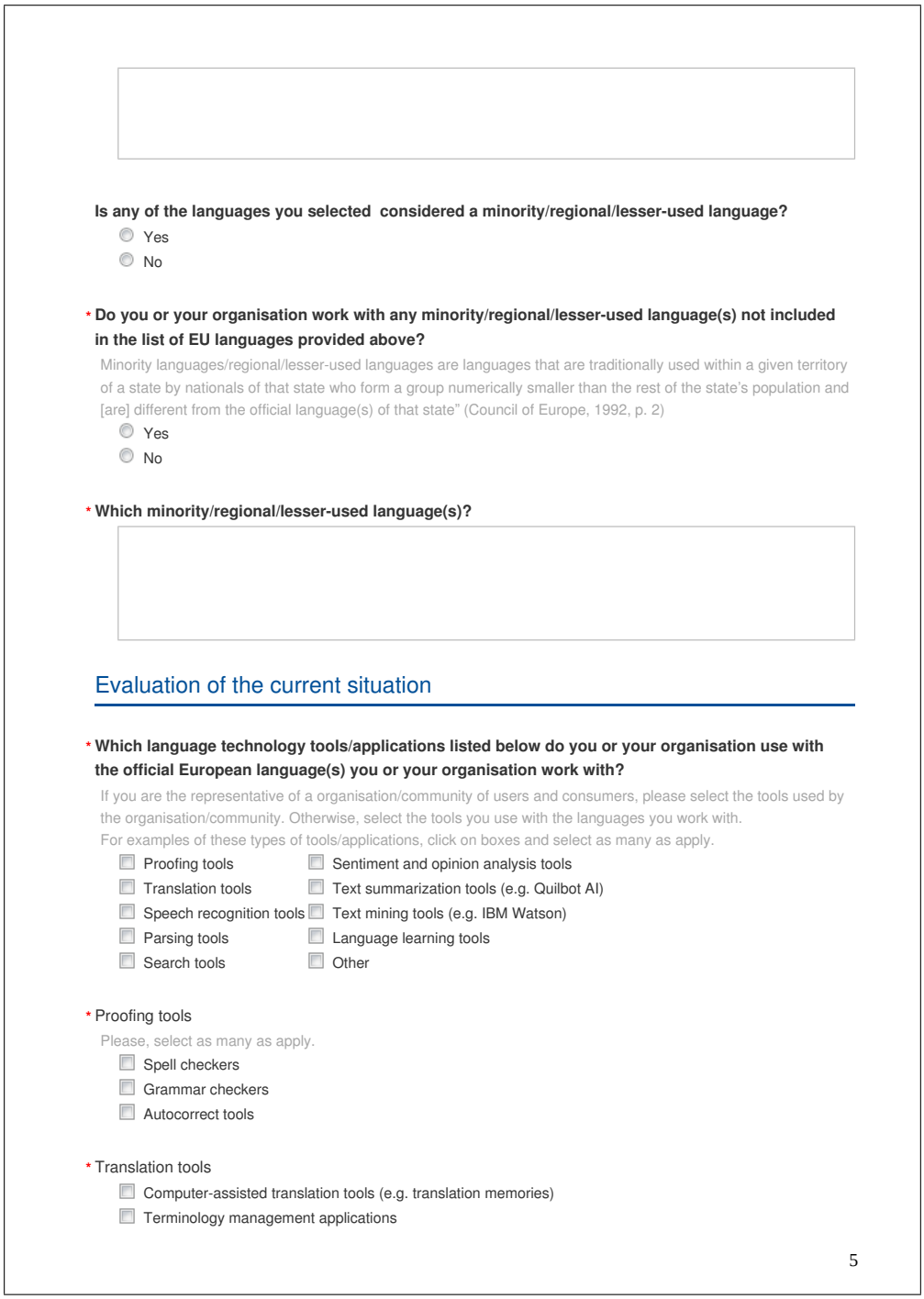
* **Which language(s)?**

- ☐ Bulgarian
- ☐ Croatian
- ☐ Czech
- ☐ Danish
- ☐ Dutch
- ☐ English
- ☐ Estonian
- ☐ Finnish
- ☐ French
- ☐ German
- ☐ Greek
- ☐ Hungarian
- ☐ Icelandic
- ☐ Irish
- ☐ Italian
- ☐ Latvian
- ☐ Lithuanian
- ☐ Maltese
- ☐ Norwegian
- ☐ Polish
- ☐ Portuguese
- ☐ Romanian
- ☐ Slovak
- ☐ Slovenian
- ☐ Spanish
- ☐ Swedish
- ☐ Other

* If "other", please specify.

4

Figure 10: Full survey as published (page 4/18)



**Is any of the languages you selected considered a minority/regional/lesser-used language?**

- Yes
- No

* **Do you or your organisation work with any minority/regional/lesser-used language(s) not included in the list of EU languages provided above?**

Minority languages/regional/lesser-used languages are languages that are traditionally used within a given territory of a state by nationals of that state who form a group numerically smaller than the rest of the state's population and [are] different from the official language(s) of that state" (Council of Europe, 1992, p. 2)

- Yes
- No

* **Which minority/regional/lesser-used language(s)?**

## Evaluation of the current situation

* **Which language technology tools/applications listed below do you or your organisation use with the official European language(s) you or your organisation work with?**

If you are the representative of a organisation/community of users and consumers, please select the tools used by the organisation/community. Otherwise, select the tools you use with the languages you work with.
For examples of these types of tools/applications, click on boxes and select as many as apply.

- Proofing tools
- Translation tools
- Speech recognition tools
- Parsing tools
- Search tools
- Sentiment and opinion analysis tools
- Text summarization tools (e.g. Quilbot AI)
- Text mining tools (e.g. IBM Watson)
- Language learning tools
- Other

* Proofing tools

Please, select as many as apply.

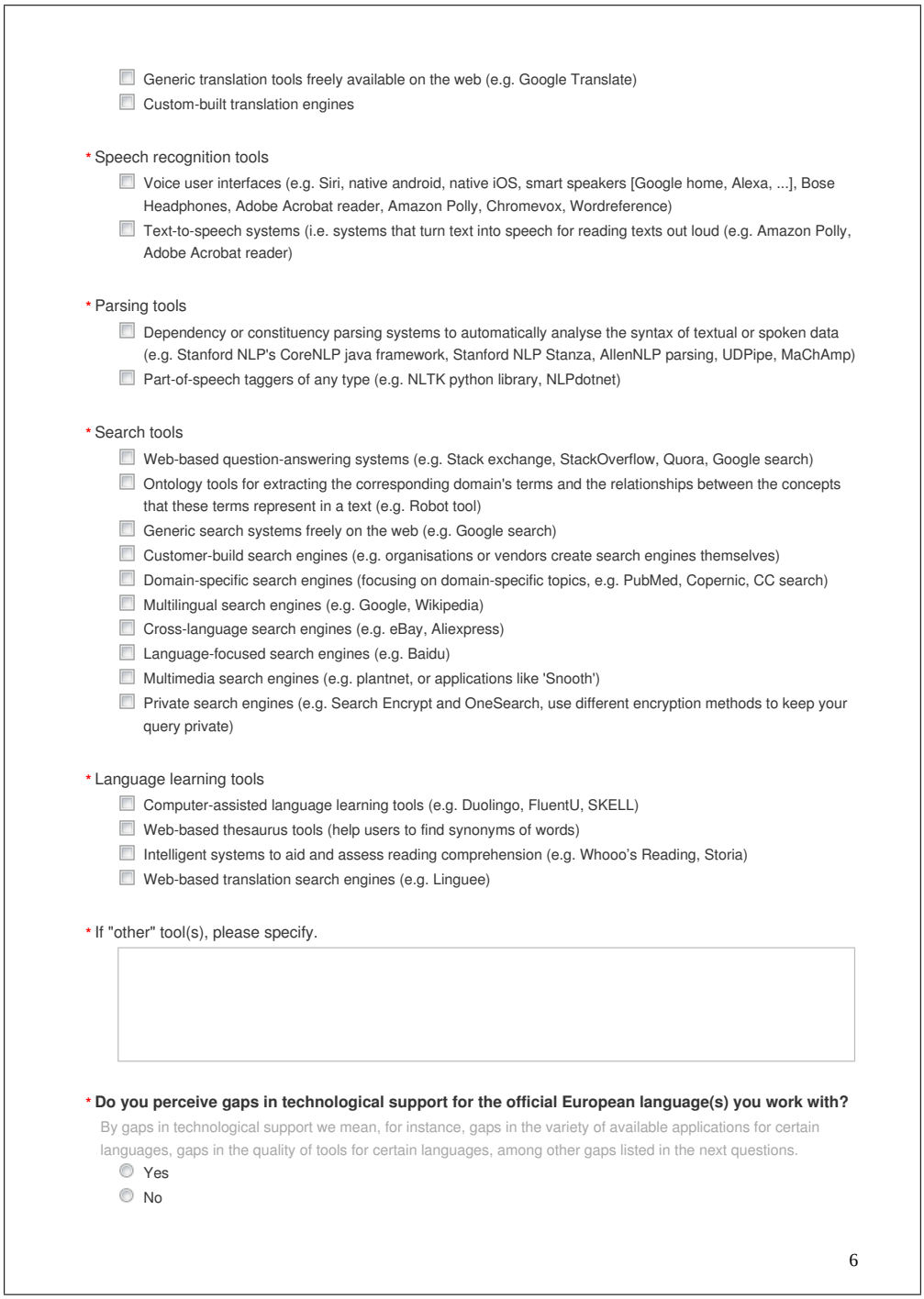
- Spell checkers
- Grammar checkers
- Autocorrect tools

* Translation tools

- Computer-assisted translation tools (e.g. translation memories)
- Terminology management applications

5

Figure 11: Full survey as published (page 5/18)



- [ ] Generic translation tools freely available on the web (e.g. Google Translate)
- [ ] Custom-built translation engines

* Speech recognition tools
  - [ ] Voice user interfaces (e.g. Siri, native android, native iOS, smart speakers [Google home, Alexa, ...], Bose Headphones, Adobe Acrobat reader, Amazon Polly, Chromevox, Wordreference)
  - [ ] Text-to-speech systems (i.e. systems that turn text into speech for reading texts out loud (e.g. Amazon Polly, Adobe Acrobat reader)

* Parsing tools
  - [ ] Dependency or constituency parsing systems to automatically analyse the syntax of textual or spoken data (e.g. Stanford NLP's CoreNLP java framework, Stanford NLP Stanza, AllenNLP parsing, UDPipe, MaChAmp)
  - [ ] Part-of-speech taggers of any type (e.g. NLTK python library, NLPdotnet)

* Search tools
  - [ ] Web-based question-answering systems (e.g. Stack exchange, StackOverflow, Quora, Google search)
  - [ ] Ontology tools for extracting the corresponding domain's terms and the relationships between the concepts that these terms represent in a text (e.g. Robot tool)
  - [ ] Generic search systems freely on the web (e.g. Google search)
  - [ ] Customer-build search engines (e.g. organisations or vendors create search engines themselves)
  - [ ] Domain-specific search engines (focusing on domain-specific topics, e.g. PubMed, Copernic, CC search)
  - [ ] Multilingual search engines (e.g. Google, Wikipedia)
  - [ ] Cross-language search engines (e.g. eBay, Aliexpress)
  - [ ] Language-focused search engines (e.g. Baidu)
  - [ ] Multimedia search engines (e.g. plantnet, or applications like 'Snooth')
  - [ ] Private search engines (e.g. Search Encrypt and OneSearch, use different encryption methods to keep your query private)

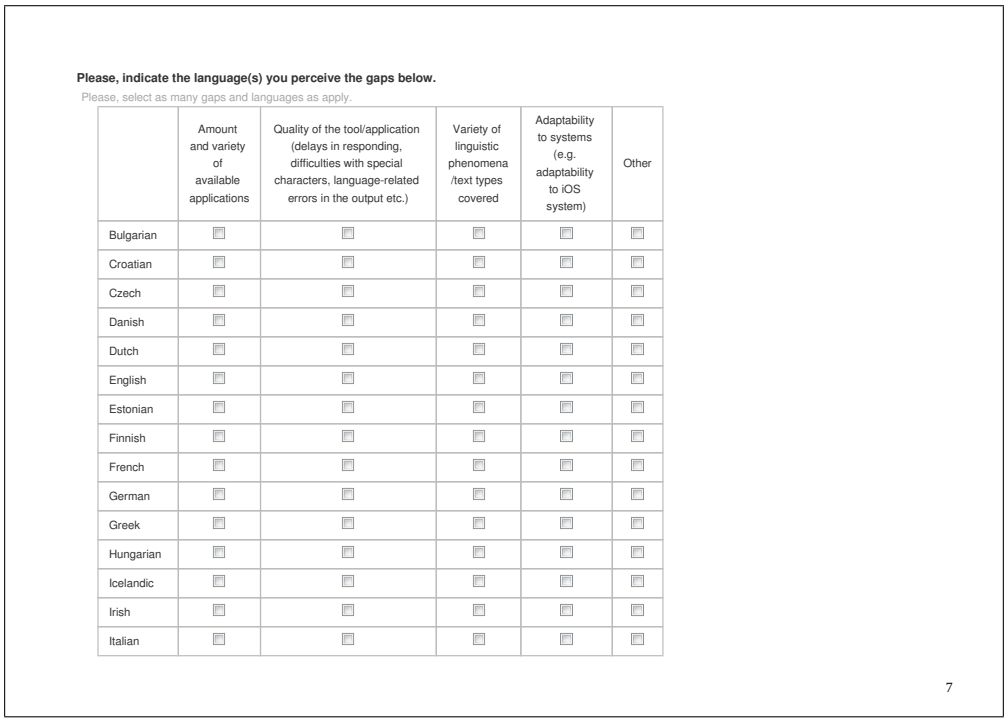
* Language learning tools
  - [ ] Computer-assisted language learning tools (e.g. Duolingo, FluentU, SKELL)
  - [ ] Web-based thesaurus tools (help users to find synonyms of words)
  - [ ] Intelligent systems to aid and assess reading comprehension (e.g. Whooo's Reading, Storia)
  - [ ] Web-based translation search engines (e.g. Linguee)

* If "other" tool(s), please specify.

* **Do you perceive gaps in technological support for the official European language(s) you work with?**
  By gaps in technological support we mean, for instance, gaps in the variety of available applications for certain languages, gaps in the quality of tools for certain languages, among other gaps listed in the next questions.
  - ( ) Yes
  - ( ) No

6

Figure 12: Full survey as published (page 6/18)



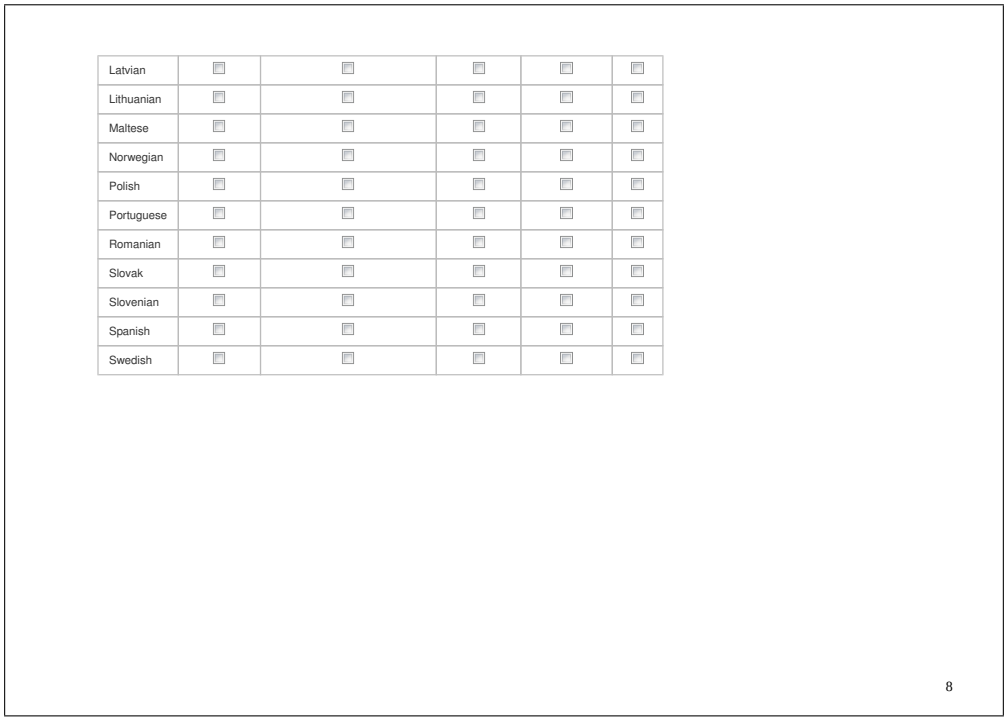
**Please, indicate the language(s) you perceive the gaps below.**

Please, select as many gaps and languages as apply.

| | Amount and variety of available applications | Quality of the tool/application (delays in responding, difficulties with special characters, language-related errors in the output etc.) | Variety of linguistic phenomena /text types covered | Adaptability to systems (e.g. adaptability to iOS system) | Other |
|---|---|---|---|---|---|
| Bulgarian | ☐ | ☐ | ☐ | ☐ | ☐ |
| Croatian | ☐ | ☐ | ☐ | ☐ | ☐ |
| Czech | ☐ | ☐ | ☐ | ☐ | ☐ |
| Danish | ☐ | ☐ | ☐ | ☐ | ☐ |
| Dutch | ☐ | ☐ | ☐ | ☐ | ☐ |
| English | ☐ | ☐ | ☐ | ☐ | ☐ |
| Estonian | ☐ | ☐ | ☐ | ☐ | ☐ |
| Finnish | ☐ | ☐ | ☐ | ☐ | ☐ |
| French | ☐ | ☐ | ☐ | ☐ | ☐ |
| German | ☐ | ☐ | ☐ | ☐ | ☐ |
| Greek | ☐ | ☐ | ☐ | ☐ | ☐ |
| Hungarian | ☐ | ☐ | ☐ | ☐ | ☐ |
| Icelandic | ☐ | ☐ | ☐ | ☐ | ☐ |
| Irish | ☐ | ☐ | ☐ | ☐ | ☐ |
| Italian | ☐ | ☐ | ☐ | ☐ | ☐ |

7

Figure 13: Full survey as published (page 7/18)



| | | | | | |
|---|---|---|---|---|---|
| Latvian | ☐ | ☐ | ☐ | ☐ | ☐ |
| Lithuanian | ☐ | ☐ | ☐ | ☐ | ☐ |
| Maltese | ☐ | ☐ | ☐ | ☐ | ☐ |
| Norwegian | ☐ | ☐ | ☐ | ☐ | ☐ |
| Polish | ☐ | ☐ | ☐ | ☐ | ☐ |
| Portuguese | ☐ | ☐ | ☐ | ☐ | ☐ |
| Romanian | ☐ | ☐ | ☐ | ☐ | ☐ |
| Slovak | ☐ | ☐ | ☐ | ☐ | ☐ |
| Slovenian | ☐ | ☐ | ☐ | ☐ | ☐ |
| Spanish | ☐ | ☐ | ☐ | ☐ | ☐ |
| Swedish | ☐ | ☐ | ☐ | ☐ | ☐ |

8

Figure 14: Full survey as published (page 8/18)

* If "other", please specify.

**In general terms, how do you evaluate the performance of the tools you use for the official European language(s) you work with?**
**Please evaluate based on a four-point scale.**

Please, evaluate as many tools as apply. If you do not know one or more tools, please select non-applicable (N/A).

| | 1.Very poor | 2. Poor | 3. Good | 4. Excellent | 5. N/A |
|---|---|---|---|---|---|
| Proofing tools (e.g. Spell checkers, Autocorrect) | | | | | |
| Translation tools (e.g. Google Translate) | | | | | |
| Speech recognition tools (e.g. Siri, Alexa) | | | | | |
| Parsing (e.g. PoS taggers) | | | | | |
| Search tools (e.g. Google search) | | | | | |
| Sentiment analysis and opinion analysis tools | | | | | |
| Text summarization (e.g. Quillbot) | | | | | |
| Text mining (e.g. IBM Watson) | | | | | |
| Language learning (e.g. Duolingo, thesaurus, bilingual dictionaries) | | | | | |
| Other | | | | | |

* If "other", please specify.

**Please choose the option that best describes the level of language technology support for the official European language(s) you or your organisation work with.**

Please, choose as many languages as apply.

| | 1. No support | 2. Poor support | 3. Good support | 4. Excellent support | 5. I do not know |
|---|---|---|---|---|---|
| Bulgarian | | | | | |
| Croatian | | | | | |
| Czech | | | | | |

9

Figure 15: Full survey as published (page 9/18)

| | | | | | |
|---|---|---|---|---|---|
| Danish | ○ | ○ | ○ | ○ | ○ |
| Dutch | ○ | ○ | ○ | ○ | ○ |
| English | ○ | ○ | ○ | ○ | ○ |
| Estonian | ○ | ○ | ○ | ○ | ○ |
| Finnish | ○ | ○ | ○ | ○ | ○ |
| French | ○ | ○ | ○ | ○ | ○ |
| German | ○ | ○ | ○ | ○ | ○ |
| Greek | ○ | ○ | ○ | ○ | ○ |
| Hungarian | ○ | ○ | ○ | ○ | ○ |
| Icelandic | ○ | ○ | ○ | ○ | ○ |
| Irish | ○ | ○ | ○ | ○ | ○ |
| Italian | ○ | ○ | ○ | ○ | ○ |
| Latvian | ○ | ○ | ○ | ○ | ○ |
| Lithuanian | ○ | ○ | ○ | ○ | ○ |
| Maltese | ○ | ○ | ○ | ○ | ○ |
| Norwegian | ○ | ○ | ○ | ○ | ○ |
| Polish | ○ | ○ | ○ | ○ | ○ |
| Portuguese | ○ | ○ | ○ | ○ | ○ |
| Romanian | ○ | ○ | ○ | ○ | ○ |
| Slovak | ○ | ○ | ○ | ○ | ○ |
| Slovenian | ○ | ○ | ○ | ○ | ○ |
| Spanish | ○ | ○ | ○ | ○ | ○ |
| Swedish | ○ | ○ | ○ | ○ | ○ |

**Please indicate based on a five-point scale how frequently you use the language technology tools /applications listed below for the languages you work with.**

Please, select as many tools as apply.

| | 1. Never | 2. Rarely | 3. Sometimes | 4. Frequently | 5. Every day |
|---|---|---|---|---|---|
| Proofing tools (e.g. Spell checkers, Autocorrect) | ○ | ○ | ○ | ○ | ○ |
| Translation tools (e.g. Google Translate) | ○ | ○ | ○ | ○ | ○ |

10

Figure 16: Full survey as published (page 10/18)

| | | | | | |
|---|---|---|---|---|---|
| Speech recognition tools (e.g. Siri, Alexa) | | | | | |
| Parsing (e.g. PoS taggers) | | | | | |
| Search tools (e.g. Google search) | | | | | |
| Sentiment analysis and opinion analysis tools | | | | | |
| Text summarization (e.g. Quillbot) | | | | | |
| Text mining (e.g. IBM Watson) | | | | | |
| Language learning (e.g. Duolingo, thesaurus, bilingual dictionaries) | | | | | |
| Other | | | | | |

* If "other" tool(s), please specify.

**Please indicate for which language(s) you or your organisation use the language technology tools /applications listed below.**

Please, select as many tools and languages as apply.

| | Proofing tools (e.g. Spell checkers, grammar checkers) | Translation tools (e.g. Google Translate) | Speech Recognition tools (e.g. Siri, Alexa) | Search tools (e.g. Google search, Wikipea) |
|---|---|---|---|---|
| Bulgarian | | | | |
| Croatian | | | | |
| Czech | | | | |
| Danish | | | | |
| Dutch | | | | |
| English | | | | |
| Estonian | | | | |
| Finnish | | | | |
| French | | | | |
| German | | | | |
| Greek | | | | |
| Hungarian | | | | |
| Icelandic | | | | |

11

Figure 17: Full survey as published (page 11/18)

| | | | | |
|---|---|---|---|---|
| Irish | ☐ | ☐ | ☐ | ☐ |
| Italian | ☐ | ☐ | ☐ | ☐ |
| Latvian | ☐ | ☐ | ☐ | ☐ |
| Lithuanian | ☐ | ☐ | ☐ | ☐ |
| Maltese | ☐ | ☐ | ☐ | ☐ |
| Norwegian | ☐ | ☐ | ☐ | ☐ |
| Polish | ☐ | ☐ | ☐ | ☐ |
| Portuguese | ☐ | ☐ | ☐ | ☐ |
| Romanian | ☐ | ☐ | ☐ | ☐ |
| Slovak | ☐ | ☐ | ☐ | ☐ |
| Slovenian | ☐ | ☐ | ☐ | ☐ |
| Spanish | ☐ | ☐ | ☐ | ☐ |
| Swedish | ☐ | ☐ | ☐ | ☐ |
| Other | ☐ | ☐ | ☐ | ☐ |

* If "other" language(s), please specify.

**Are there language technology tools/applications available for the minority/regional/lesser-used language(s) you or your organisation work with?**

- ○ Yes
- ○ No
- ○ I do not know

* **Which tools/applications do you use with these minority/regional/lesser-used languages?**

For more examples of these types of tools, click on the boxes and select as many tools as apply.

- ☐ Proofing tools
- ☐ Translation tools
- ☐ Speech recognition tools
- ☐ Parsing tools
- ☐ Search tools
- ☐ Sentiment and opinion analysis tools
- ☐ Text summarization tools (e.g. Quilbot AI)
- ☐ Text mining tools (e.g. IBM Watson)
- ☐ Language learning tools
- ☐ Other

* Proofing tools

Select as many as apply.

- ☐ Spell checkers
- ☐ Grammar checkers
- ☐ Autocorrect

12

Figure 18: Full survey as published (page 12/18)

* Translation tools

Select as many as apply.

- Computer-assisted translation tools (e.g. translation memories)
- Terminology management applications
- Generic translation tools freely available on the web (e.g. Google Translate)
- Custom-built translation engines

* Speech recognition/synthesis tools

Select as many as apply.

- Voice user interfaces (e.g. Siri, native android, native iOS, smart speakers [Google home, Alexa, ...], Bose Headphones, Adobe Acrobat reader, Amazon Polly, Chromevox, Wordreference)
- Text-to-speech systems (i.e. systems that turn text into speech or for reading text out loud (e.g. Amazon Polly, Adobe Acrobat reader)

* Parsing tools

Please, select as many as apply.

- Dependency or constituency parsing systems to automatically analyse the syntax of textual or spoken data (e.g. Stanford NLP's CoreNLP java framework, Stanford NLP Stanza, AllenNLP parsing, UDPipe, MaChAmp)
- Part-of-speech taggers of any type (e.g. NLTK python library, NLPdotnet)

* Search tools

Please, select as many as apply.

- Web-based question-answering systems (e.g Stack exchange, StackOverflow, Quora, Google search)
- Ontology tools for extracting the corresponding domain's terms and the relationships between the concepts that these terms represent in a corpus (e.g. Robot tool)
- Generic search systems freely on the web (e.g. Google search)
- Customer-build search engines (e.g organisations or vendors create search engines themselves)
- Domain-specific search engines (focusing on domain-specific topics, e.g. PubMed, Copernic, CC search)
- Multilingual search engines (e.g. Google, Wikipedia)
- Cross-language search engines (e.g. eBay, Aliexpress)
- Language-focused search engines (e.g. Baidu)
- Multimedia search engines (e.g. plantnet, or applications like 'Snooth')
- Private search engines (e.g. Search Encrypt and OneSearch, use different encryption methods to keep your query private)

* Language learning tools

Please, select as many as apply.

- Computer-assisted language learning tools (e.g. Duolingo, FluentU, SKELL)
- Web-based thesaurus tools (help users to find synonyms of words e.g. thesaurus.com)
- Intelligent systems to aid and assess reading comprehension (e.g. Whooo's Reading, Storia)
- Web-based translation search engines (e.g. Linguee)

* If "other", please specify.

13

Figure 19: Full survey as published (page 13/18)

**Do you perceive gaps in technological support for the minority/regional/lesser-used language(s) you work with?**

By gaps in technological support we mean, for instance, gaps in the variety of available applications for certain languages, gaps in the quality of tools for certain languages, among other gaps listed in the next questions.

- ○ Yes
- ○ No

**Please, indicate the gap(s) you perceive.**

Please, select as many as apply.

- ☐ Gaps in the amount and variety of available applications
- ☐ Gaps in the quality of the tool/application (delays in responding, difficulties with special characters, language-related errors in the output etc.)
- ☐ Gaps in the variety of linguistic phenomena/text types covered
- ☐ Gaps in adaptability to systems (e.g. adaptability to iOS system)
- ☐ Not sure
- ☐ Other

* If "other", please specify.

**In general terms, how do you evaluate the performance of the language technology tools for the minority/regional/lesser-used language(s) you work with? Please evaluate based on a four-point scale.**

Please, select as many tools as apply. If you cannot evaluate for any reason, please select not applicable (N/A).

| | 1.Very poor | 2. Poor | 3. Good | 4. Excellent | 5. N/A |
|---|---|---|---|---|---|
| Proofing tools (e.g. Spell checkers, Autocorrect) | ○ | ○ | ○ | ○ | ○ |
| Translation tools (e.g. Google Translate) | ○ | ○ | ○ | ○ | ○ |
| Speech recognition tools (e.g. Siri, Alexa) | ○ | ○ | ○ | ○ | ○ |
| Parsing (e.g. PoS taggers) | ○ | ○ | ○ | ○ | ○ |
| Search tools (e.g. Google search) | ○ | ○ | ○ | ○ | ○ |
| Sentiment analysis and opinion analysis tools | ○ | ○ | ○ | ○ | ○ |

14

Figure 20: Full survey as published (page 14/18)

| | | | | | |
|---|---|---|---|---|---|
| Text summarization (e.g. Quillbot) | | | | | |
| Text mining (e.g. IBM Watson) | | | | | |
| Language learning (e.g. Duolingo, thesaurus, bilingual dictionaries) | | | | | |
| Other | | | | | |

If "other", please specify.

**Please, choose the option that best describes the level of language technology support for the minority/regional/lesser-used language(s) you or your organisation work with.**

Please, select as many tools as apply. If you do not know one or more tools, select not applicable (N/A).

| | 1. Very poor | 2. Poor | 3. Good | 4. Excellent | 5. N/A |
|---|---|---|---|---|---|
| Proofing tools (e.g. Spell checkers, Autocorrect) | | | | | |
| Translation tools (e.g. Google Translate) | | | | | |
| Speech recognition tools (e.g. Siri, Alexa) | | | | | |
| Parsing (e.g. PoS taggers) | | | | | |
| Search tools (e.g. Google search) | | | | | |
| Sentiment analysis and opinion analysis tools | | | | | |
| Text summarization (e.g. Quillbot) | | | | | |
| Text mining (e.g. IBM Watson) | | | | | |
| Language learning (e.g. Duolingo, thesaurus, bilingual dictionaries) | | | | | |
| Other | | | | | |

* If "other", please specify.

15

Figure 21: Full survey as published (page 15/18)

**Please indicate based on a five-point scale how frequently you use the language technology tools /applications listed below for the minority/regional/lesser-used languages you work with.**

Please, select as many tools as apply.

| | 1. Never | 2. Rarely | 3. Sometimes | 4. Frequently | 5. Every day |
|---|---|---|---|---|---|
| Proofing tools (e.g. Spell checkers, Autocorrect) | ○ | ○ | ○ | ○ | ○ |
| Translation tools (e.g. Google Translate) | ○ | ○ | ○ | ○ | ○ |
| Speech recognition tools (e.g. Siri, Alexa) | ○ | ○ | ○ | ○ | ○ |
| Parsing (e.g. PoS taggers) | ○ | ○ | ○ | ○ | ○ |
| Search tools (e.g. Google search) | ○ | ○ | ○ | ○ | ○ |
| Sentiment analysis and opinion analysis tools | ○ | ○ | ○ | ○ | ○ |
| Text summarization (e.g. Quillbot) | ○ | ○ | ○ | ○ | ○ |
| Text mining (e.g. IBM Watson) | ○ | ○ | ○ | ○ | ○ |
| Language learning (e.g. Duolingo, thesaurus, bilingual dictionaries) | ○ | ○ | ○ | ○ | ○ |
| Other | ○ | ○ | ○ | ○ | ○ |

* If "other" tool, please specify.

## Predictions and visions for future

* **In your opinion, what provision of resources would increase the use of language tools for the specific languages you or your organisation use?**

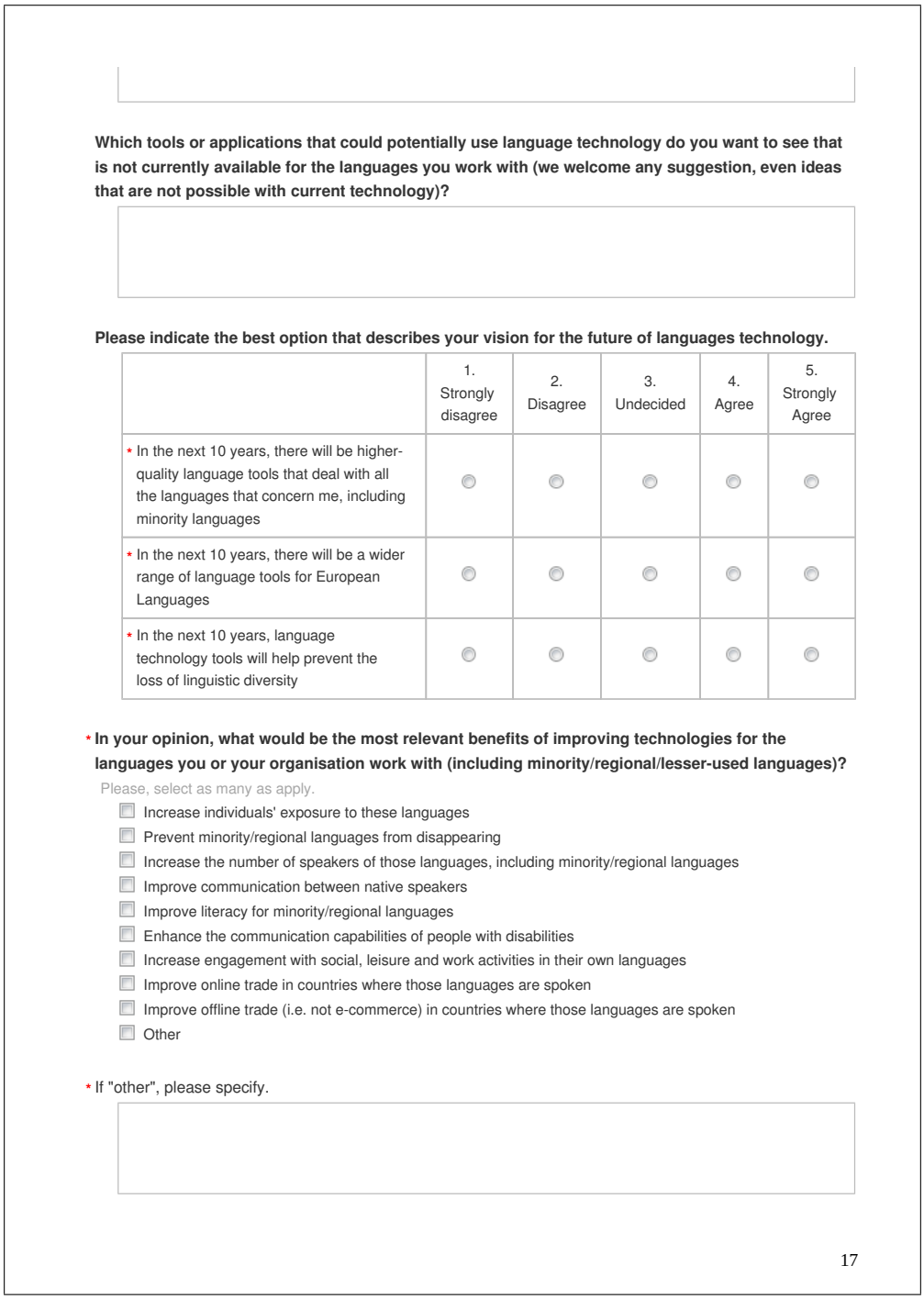
Please, select as many as apply.

- ☐ A wider range of language tools for the languages I work with
- ☐ Higher-quality tools for the languages I work with
- ☐ More training of personnel dealing with such tools
- ☐ Other

* If "other", please specify.

16

Figure 22: Full survey as published (page 16/18)



**Which tools or applications that could potentially use language technology do you want to see that is not currently available for the languages you work with (we welcome any suggestion, even ideas that are not possible with current technology)?**

**Please indicate the best option that describes your vision for the future of languages technology.**

| | 1. Strongly disagree | 2. Disagree | 3. Undecided | 4. Agree | 5. Strongly Agree |
|---|---|---|---|---|---|
| * In the next 10 years, there will be higher-quality language tools that deal with all the languages that concern me, including minority languages | | | | | |
| * In the next 10 years, there will be a wider range of language tools for European Languages | | | | | |
| * In the next 10 years, language technology tools will help prevent the loss of linguistic diversity | | | | | |

* **In your opinion, what would be the most relevant benefits of improving technologies for the languages you or your organisation work with (including minority/regional/lesser-used languages)?**

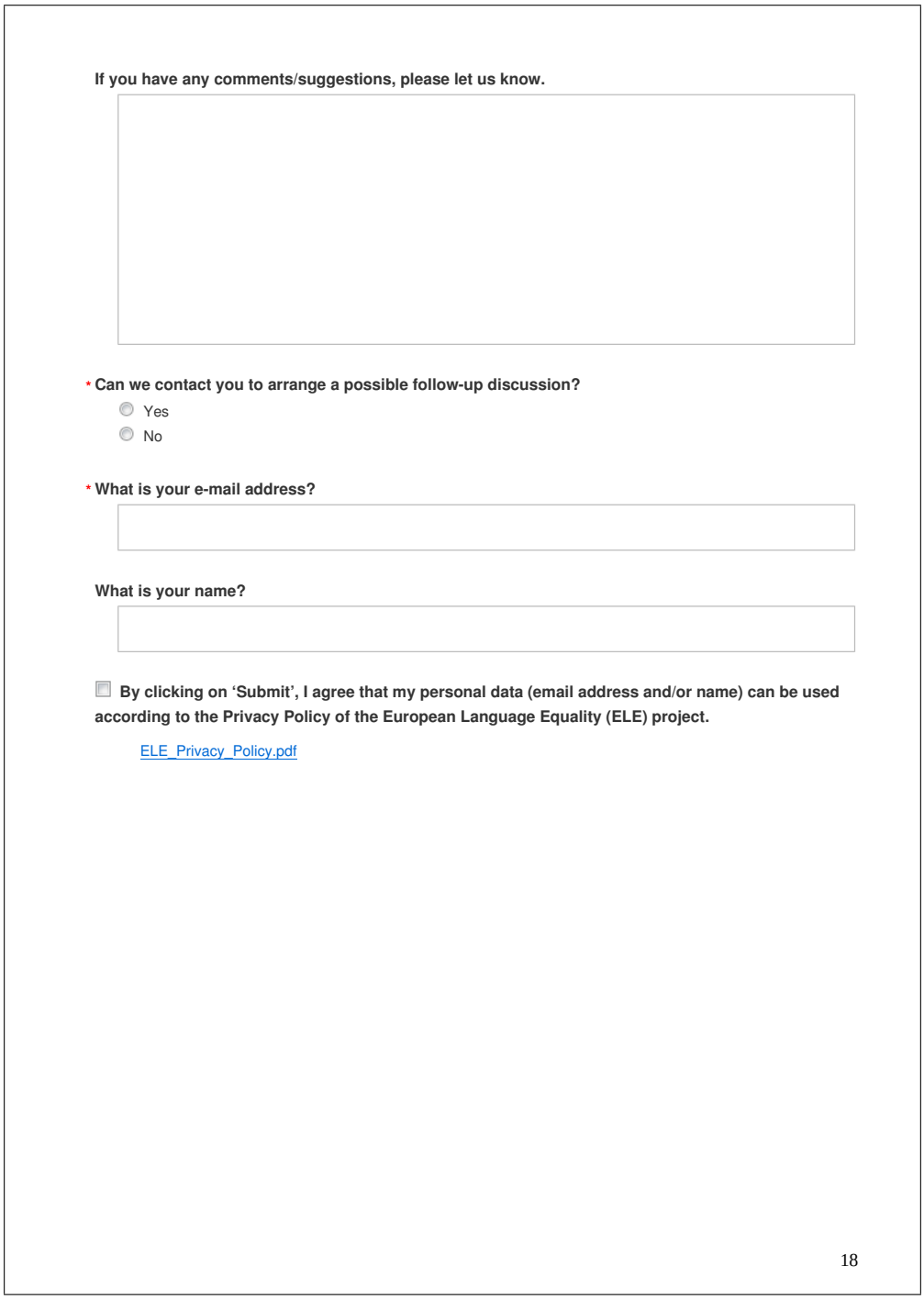
Please, select as many as apply.

- Increase individuals' exposure to these languages
- Prevent minority/regional languages from disappearing
- Increase the number of speakers of those languages, including minority/regional languages
- Improve communication between native speakers
- Improve literacy for minority/regional languages
- Enhance the communication capabilities of people with disabilities
- Increase engagement with social, leisure and work activities in their own languages
- Improve online trade in countries where those languages are spoken
- Improve offline trade (i.e. not e-commerce) in countries where those languages are spoken
- Other

* If "other", please specify.

17

Figure 23: Full survey as published (page 17/18)



If you have any comments/suggestions, please let us know.

* Can we contact you to arrange a possible follow-up discussion?

- Yes
- No

* What is your e-mail address?

What is your name?

By clicking on 'Submit', I agree that my personal data (email address and/or name) can be used according to the Privacy Policy of the European Language Equality (ELE) project.

ELE_Privacy_Policy.pdf

18

Figure 24: Full survey as published (page 18/18)

# B. Additional tables and graphs

| Country | Answers count | % |
|---|---|---|
| Spain | 7 | 26.9 |
| France | 5 | 19.2 |
| Germany | 4 | 15.4 |
| Belgium | 3 | 11.5 |
| Portugal | 2 | 7.7 |
| Croatia | 1 | 3.8 |
| Greece | 1 | 3.8 |
| Norway | 1 | 3.8 |
| Italy | 1 | 3.8 |
| USA | 1 | 3.8 |

Table 2: Breakdown of answers count to questions “In which country are you based? if “other”, please speficy”

| Types of organisations | Answers count | % |
|---|---|---|
| Education/research | 12 | 46,2 |
| large enterprises | 6 | 23,1 |
| SME | 3 | 11,5 |
| Innovation | 2 | 7,7 |
| Professional association | 2 | 7,7 |
| Independent contractor/ consultant | 1 | 3,8 |

Table 3: Breakdown of answers count to the question “Which of the following best describes the type of organisation you work for?”

| Languages | Answers count | % |
|---|---|---|
| English | 25 | 96.2 |
| Spanish | 13 | 50 |
| French | 11 | 42.3 |
| German | 10 | 38.5 |
| Italian | 7 | 26.9 |
| Portuguese | 5 | 19.2 |
| Greek | 3 | 11.5 |
| Dutch | 3 | 11.5 |
| Catalan | 2 | 7.7 |
| Croatian | 2 | 7.7 |
| Japanese/chinese | 1 | 3.8 |
| Basque | 1 | 3.8 |
| Slovenian | 1 | 3.8 |
| Norwegian | 1 | 3.8 |
| Czech | 1 | 3.8 |

Table 4: Breakdown of answers to the question “Which of the official European language(s) listed below do you or your organisation work with? if “other”, please speficy”

| Language Technologies | Answers count | % |
|---|---|---|
| **Proofing tools** | | |
| Spell checkers | 18 | 69.2 |
| Grammar Checkers | 18 | 69.2 |
| Autocorrect tools | 19 | 73.1 |
| **Translation tools** | | |
| Generic translation tools freely available on the web | 19 | 73.1 |
| Computer-assisted translation tools | 8 | 30.8 |
| Terminology management applications | 3 | 11.5 |
| Custom-built translation engines | 2 | 7.7 |
| **Speech recognition tools** | | |
| Voice user interfaces | 5 | 19.2 |
| Text-to-speech systems | 6 | 23.1 |
| **Parsing tools** | | |
| Dependency or constituency parsing | 2 | 7.7 |
| Part-of-speech taggers of any type | 0 | 0 |
| **Search tools** | | |
| Generic search systems freely on the web | 19 | 73.1 |
| Multilingual search engines | 18 | 69.2 |
| Web-based question-answering systems | 9 | 34.6 |
| Cross-language search engines | 4 | 15.4 |
| Domain-specific search engines | 3 | 11.5 |
| Customer-build search engines | 3 | 11.5 |
| Private search engines | 2 | 7.7 |
| **Language Learning tools** | | |
| Ontology tools | 2 | 7.7 |
| Web-based translation search engines | 11 | 42.3 |
| Web-based thesaurus tools | 6 | 23.1 |
| Computer-assisted language learning tools | 4 | 15.4 |
| Intelligent systems to aid and assess reading comprehension | 0 | 0 |

Table 5: Breakdown of answers to the question: “Which language technology tools or applications listed below do you or your organisation use with the official European language(s) you or your organisation work with? if “other”, please specify”