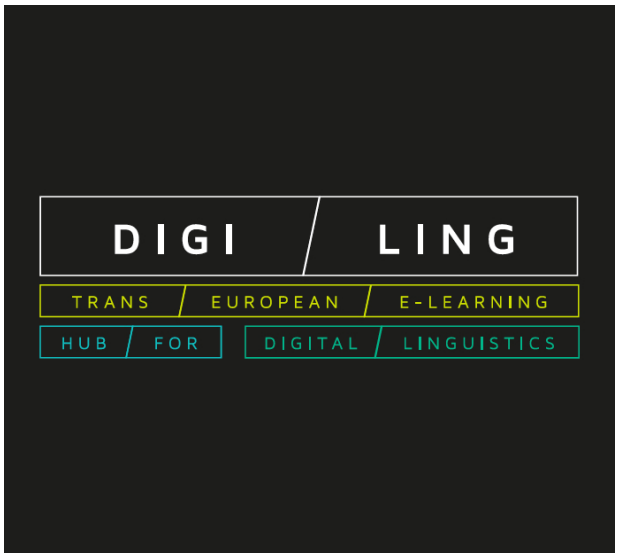


Project title	DigiLing: TransEuropean e-Learning Hub for Digital Linguistics
Deliverable title	**Labour market needs survey and the DigiLing model curriculum**
Authors	Špela Vintar, Ondřej Matuška, Jure Škerl, Petra Bago, Dragos Ciobanu, Vaclav Cvrček, Oliver Čulo, Miloš Jakubiček, Nives Mikelić Preradović, Jean Nitzke, Alina Secara, Zdenek Žabokrtsky
Date	28 April 2017
Version	v1
Dissemination level	public

This document is part of the KA2 Strategic Partnership DigiLing: TransEuropean e-Learning Hub for Digital Linguistics.
This project has received funding from the European Union's Erasmus+ Programme under grant agreement no. 16-203-021558.

Contents

1 Introduction

The DigiLing: Trans-European e-Learning Hub for Digital Linguistics project is aimed at creating new online courses for Digital Linguistics, a new interdisciplinary field of study at the crossroads between linguistics, information technologies and social sciences. Digital Linguistics is not synonymous to Computational Linguistics (CL), although certain skills and research methods may overlap between these two disciplines. While CL is primarily concerned with modelling and processing language in computer applications in order to provide language-aware tools and systems, the emerging field of Digital Linguistics is broader in the sense that it should provide the complete set of scientific, methodological and practical foundations pertaining to communication in the digital age.

This subsumes linguistic knowledge, such as native and foreign language competence, translation-related competencies and an understanding of language analysis procedures at all levels, and it also entails Natural Language Processing skills, particularly at the level of an in-depth understanding of state-of-the-art NLP techniques and basic programming skills. But on top of this "traditional" interdisciplinary blend between linguistics and IT, Digital Linguistics has several additional foci which justify its claim for a field of study in its own right.

The first is digital content authoring, which could at first glance be understood as one of the linguistic competencies formerly known as text production skills developed through text and discourse studies. But the digital age has brought profound changes to the ways texts and other types of content are produced. Digital news media have revolutionized journalism and brought new paradigms into the concepts of journalistic research, credibility, authenticity, authorship and accessibility. The personalization of digital services means that content, including web sites, ads, user manuals and posts on social platforms, is produced in a targeted and user-centred fashion, whereby the cyber-identity of the target user is not to be confused with their real-world identity.

These issues can only be adequately addressed by bringing in the sociological, psychological and cognitive perspectives, and by putting communicative behaviour in digital media into the centre of study. Another aspect of content authoring is related to multilingual contents and activities such as translation, localization, subtitling and interpreting. While traditionally the providers of multilingual services were the ones generating content, contemporary translators compose texts by selecting from available hits offered by translation memories, machine translation engines and other multilingual resources. From the cognitive point of view, as Pym (2013) points out, the process of [content] generation has been transformed into the process of selection, where the issue of critical assessment and trust has become paramount.

The issues of trust, identity, authorship and reuse inevitably lead to questions concerning intellectual property rights and data protection, but also ethical aspects of communication

in digital media. The legislative framework which attempts to regulate rights related to language data is lagging behind – for example the EU Copyright Directive dates back to 2001 and is no longer applicable to many scenarios we encounter today. Significant steps to better define the legal framework around language resources, their distribution and reuse were made within the QTLaunchPad project (Tsavos et al. 2014), see also TAUS[1] (2013).

Therefore, Digital Linguistics as a field of study combines insights and perspectives from different disciplines and does not overlap with Computational Linguistics, nor for that matter with Digital Humanities, Sociolinguistics or Corpus Linguistics, though it may inherit methods and tools from all of the above.

In the remainder of this report we justify the need for DL by giving an overview of the relevant market surveys and analyses of the global language industry, then continue to describe the designing and dissemination of the DigiLing survey. The responses to our survey serve as the basis for the model curriculum in Digital Linguistics which is the main outcome of this project activity.

2 Overview of existing surveys

In the following section we present an overview of recent language industry surveys which give important insights into the trends and challenges in the fast-growing domain of language services.

The main sources considered are the following:

- 2016 Language Industry Survey: Expectations and Concerns of the European Language Industry. The survey was performed jointly by ELIA/EMT/EUATEC/GALA/LIND (LIS16) and has received 445 responses from 35 countries.
- TAUS Translation Technology Report 2016 (TAUS16); this document is composed every few years by the language industry think tank, the Translation Automation User Society. Formerly available on TAUS website, as of 2016 the report is only available to TAUS members.
- SDL Language Technology Insights 2016 (SDL16) represents a strategic document analyzing the trends as perceived by one of the largest language service and technologies provider SDL. Available to registered users in the form of an executive summary.
- Gaspari et al. 2015: A survey of machine translation competences: Insights for translation technology educators and practitioners. Perspectives, 23:3, 333-358 (Gaspari et al. 2015). The paper reports on the findings of a survey on MT which collected 438 responses.

[1] https://www.taus.net/think-tank/articles/translate-articles/clarifying-copyright-on-translation-data

- Schmitt, Peter A., Lina Gerstmeyer, and Sarah Müller. *Übersetzer und Dolmetscher: Eine internationale Umfrage zur Berufspraxis*. 2016. The book is a joint prublication and comprises different national and international surveys on the translator's and interpreter's job market

On the general scale, the language industry still seems to be one of the fastest growing industries in the world. The growth of the worldwide language services market between 2014 and 2016 was 5.52%, with the size of the global language industry estimated at $40 Billion. The projected growth rate in the period 2017-2018 lies between 6.5 and 7.5% with an estimated turnaround for 2020 of $45 Billion (source: Common Sense Advisory/GALA).

LIS16, which focuses on European companies, measures the optimism amongst LSPs by comparing the number of companies planning expansion to the number of companies planning to close down or move their business. It seems that in 2016 the sentiment amongst LSPs was less optimistic than in 2015 (3.19 -> 2.5), however when the size of companies was considered still very positive with larger LSPs planning more expansion.

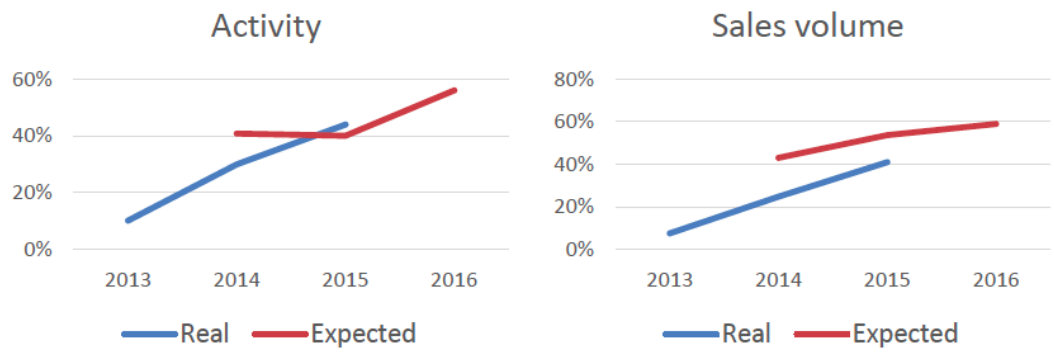


Figure 1: Activity and sales levels of European LSPs (source: LIS16)

Figures 1 and 2 show the real and expected trends in the activity, sales volume and rates. A negative trend is observable with regard to pricing. Rates charged for language services have been continually falling since 2014 and this trend is not expected to change anytime soon.

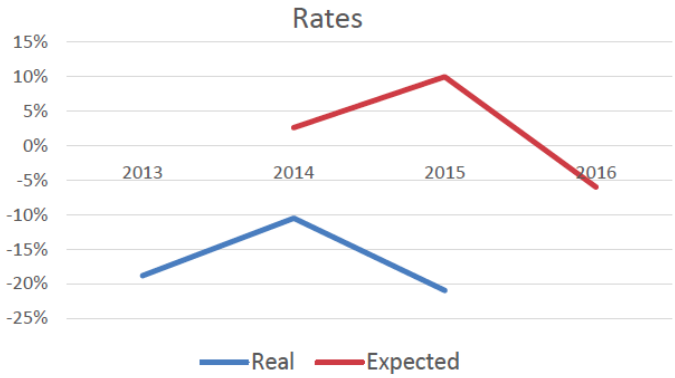


Figure 2: Rates of services (source: LIS16)

TAUS in its technology report predicts five main trends and challenges which will govern the field of language and translation services in near future. The first of these trends is the so-called datafication of translation, referring to the increasingly varied range of data sources which are instantly accessible during translation and relieve the translator of much of the information mining work. Translation memories, various MT engines, term bases, autosuggest dictionaries and other resources provide hits from which the translator has to choose from and employ primarily target language competence to produce functional content. From the cognitive point of view, the primary unit of translation has shifted from text through segment to reusable fragment, and the process of (text) generation has been transformed into a process of (reusable fragment) selection and reordering (Pym 2013).

The second challenge is the marriage between Machine Translation and Artificial Intelligence, referring to fast technological advances in the form of neural MT and sophisticated Machine Learning techniques which are likely to make MT more human-like. The third pertains to quality evaluation and assurance – the language industry will need to implement more systematic instruments of quality monitoring into its workflows and make it into an industry benchmark.

The fourth trend is the power of crowds. Crowdsourced translations used to be the landmark of enthusiastic communities such as Facebook users, open source software projects or movie fans contributing free subtitles. However, in the past few years platforms like Ackuna function as intermediaries between clients and crowds, thus reselling services which have been performed for free. The last challenge identified by TAUS is speech-to-speech translation, as more and more content in the digital world is in spoken format.

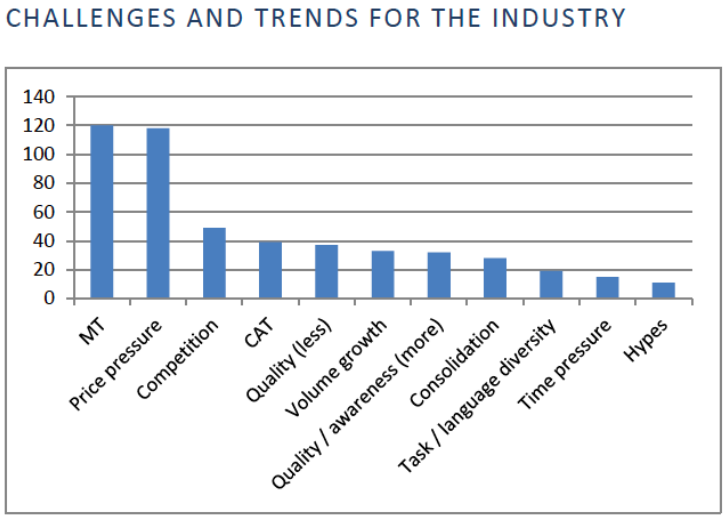


Figure 3: Challenges and trends for the industry (source: LIS16)

The LIS16 survey identifies similar trends and challenges perceived by the industry (Figure 3), with Machine Translation and price pressure far ahead of other issues such as competition and quality. Figure 4 displays an even more relevant list of challenges in the context of this report as it compares results from previous surveys and reveals the topics which appear in 2016 for the first time. Among them we find a range of challenges which

are very much related to the skills and competencies of the (potential) employees, e.g. the skills gap, new services, staff recruitment and how to differentiate.

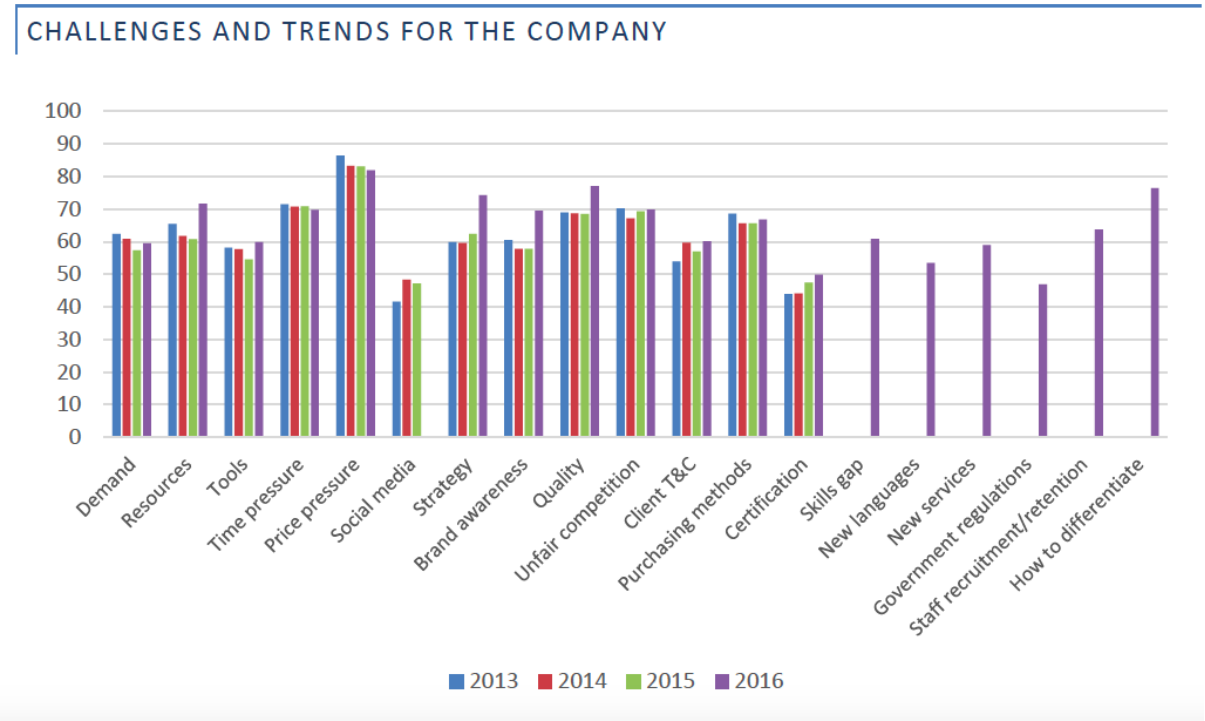


Figure 4: Challenges and trends for the company (Source: LIS16)

Machine translation is perceived both as a major opportunity, a trend and a challenge. It is therefore interesting that according to several available surveys the majority of LSPs still use no MT (LIS16: 59%, SDL16: 60%, Gaspari et al.: 58%). We should however take into account that these percentages are calculated from the total number of respondents regardless of their size. A more detailed breakdown of responses by company size is given in LIS16 and shows a clearly discernible difference between large LSPs, of which only 29% use no MT, and smaller companies which predominantly work with CAT tools and publicly available MT engines, if at all. Still, according to TAUS16 the future lies very much in the realm of automated translation:

> *What we predict is that the world will get accustomed to what we call Fully Automatic Useful Translation (FAUT) and will more and more accept this as the norm for standard translation.*

In view of these trends - growing automation and falling rates – quality becomes a widely acknowledged concern. The SDL16 technology insights reveal that clients consider quality more important than cost or speed, and similar findings are reported in the Gaspari15 survey. Apparently many LSPs still use no systematic quality control mechanisms and rework is common. The prevailing cause of the clients' dissatisfaction is inconsistent terminology, yet many LSPs employ no systematic terminology management solutions.

Schmitt, Gerstmayer and Müller (2016) concentrate on the job perspectives and needs of professional translators and interpreters. According to Schmitt et al. (ibid.: 24), professional translators still occasionally consider translation theory as irrelevant. On the other hand,

universities reject to teach only practical courses that prepare translator's solely to fit economy. He reports from different surveys that machine translation plays a vital role in freelancers' work environment (ibid.: 42). This was also confirmed by a myGengo study that found in 2009 that the three most influential trends are machine translation, hybrid translation (combing MT and post-editing), and community- and crowd-translations (2016: 51). Only 10% of the CUITI members, however, use MT (ibid.: 71). The three sectors that need language services the most are related technology, medicine, and advertisement/marketing. The top three fastest growing languages services and technologies have been translation, web globalisation and software localisation (in this order) in 2010-2013. Multimedia localisation has been on place five or six and translation technologies on place five to seven (ibid.: 46-47). Most translators that are members at CUITI[2] are freelancers (64%) and (only) 75% use CAT-tools, where SDL Trados studio is by far the tool used most often (ibid.: 66-70). Similarly, most participants use SDL Multiterm as a terminology management tool (ibid.: 179). When asked for what the participants considered needs to be more integrated into the curricula, the most wished for option was learning to use a PC and CAT-tools properly.

3 DigiLing labour market needs survey

The aim of the survey was to assess the needs of various types of employers from the language industry as well as non-language industry for expertise in the area of language-related tasks. The questionnaire was designed to provide insight into the ways companies deal with (unstructured) textual data, such as emails, business correspondence, documentation, websites, support tickets etc., but also audiovisual types of content. Inevitably, this lead to including topics and concepts that especially the general public might not be familiar with, such as text categorization, domain modeling, term extraction, localization and similar. On the one hand we wanted the survey to be as comprehensive as possible and include a wide variety of possible tasks from web content authoring to multimodal interfaces, while on the other hand it was essential to keep the questionnaire easy to understand and manageable. To achieve this, a two-stage approach to the survey was adopted.

3.1 The pilot stage

The aim of the pilot survey was to test whether respondents will be able to understand and provide relevant responses when asked about concepts and terminology they may not be familiar with. This was tested in face-to-face interviews. A pilot survey form was designed to provide a structure to the interview but the respondents were encouraged to also give informal unstructured information whenever they felt like providing additional information.

2 The survey is presented in detail by Gerstmayer in the book.

The interviewers also took notes of any questions which indicated potentially unclear wording or inadequate use of terminology.

It turned up during the pilot survey that there are a few spots with ambiguous wording and two questions with different wording to ask for virtually the same information.
It also became obvious that *automated communication technology* is generally looked down on but after providing more details about possible deployment of such technology, the respondents became interested and could think of use cases in their companies. This was taken into consideration for the final survey and the questions were phrased as more explanatory. In addition, the topic of human and machine translation proved to be more complex and was redesigned into a more complex section in the final survey.

The survey was designed as a simple Google Docs documents so that interviewers could share their views on each others comments. The pilot survey was monolingual, in English.

3.2 Performing the pilot interviews

3.2.1 Slovenia

In Slovenia, the pilot stage was conducted by setting up interviews with two local companies.

The first one was Dnevnik, d.d., a newspaper company which publishes one of the largest dailies in Slovenia. The interview was conducted in person and lasted around 45 minutes. Its purpose was mainly to ascertain whether the questions were clear enough even to a respondent with no background in IT or digital linguistics. Whenever a question seemed to elicit confusion an oral explanation was given as to what information precisely was sought. The interviewee was not an IT expert and certain terminology needed to be explained to them. With regard to the questions themselves, the respondent considered them clear enough.

A critical point that came up during the interview was the inconsistent use of the terms “need” and “use”. Some questions asked whether there existed a need for a certain language technology solution while others asked whether a solution was already being used. Based on this, the Ljubljana team suggested that the questions be made more consistent in terms of what information they were asking for. In the final version of the survey, all questions were changed so that they asked specifically about the existence of needs for language technologies and not their actual implementation.

A small insight was also gained about the outsourcing of the tasks and the multilinguality. None of the tasks mentioned by the pilot survey were being outsourced and there was no multilinguality in the company. The respondent thought that asking about these two was unnecessary. As it turned out in the final survey, this was the case in a considerable number of companies, particularly in smaller ones, but certainly not in all of them. Therefore the

decision to disregard this remark and retain questions about outsourcing and multilinguality turned out to be correct.

The second pilot interview was performed with the CEO of EBA d.o.o., a small company providing intelligent customized document management solutions to clients. As it turned out this company would be in immediate need of digital linguists as they already perform most of the tasks our survey was targeting, however not in a very sophisticated manner regarding language processing. EBA d.o.o. serves large clients and companies both from the private and public sector, but since their team of developers is composed exclusively of IT experts they seek to improve their main product with regard to the language technologies used in the processing of documents and other textual data. The main feedback on the pilot survey concerned an occasional lack of clarity in the distinction between the responding company and the clients it serves. For example, the survey inquired about the need of analyzing competitors' websites, and if the respondent is a company to which large clients outsource linguistic tasks it becomes unclear whether they should respond in their own name or that of their clients.

3.2.2 UK

Leeds Team conducted the Pilot Survey with the CEO of Andiamo! Language Services Ltd (http://www.andiamo.co.uk/), a medium-size professional language services provider based in Leeds. Andiamo! offers localisation services specialising in the automotive, technical and engineering domains.

The specific results to the pilot questionnaire were shared with the team and included in the Pilot Survey _ summary of responses report. Generally, the results and comments gathered allowed us to address several issues:

- Modify questions to enable more companies to reply to survey by avoiding the impression that it is aimed at very large companies only;
- Acknowledge language industry expectations regarding linguists and non-linguistic tasks (automation) and the role of linguists in their completion;
- Inclusion of a No answer/N/A option in the final survey.

3.2.3 Germany

To test the survey in Germany, an interview was performed with an employee of an IT company specialising on public transportation. The respondent is employed as a software developer. Due to a former training in translation, our respondent was especially aware of language use and needs in the company. During the interview, they suggested that an additional option of "no" would be necessary in the questions that only contained the answer options "already doing", "planning to" and "not sure", because they were certain that some of the suggested needs were not and would not be implemented in the company. Further, they commented on the wording of some phrases like "Google-like technologies".

3.2.4 Croatia

The Zagreb Team conducted the interview for the pilot stage of the survey with two employees of an energy, utilities & mining company, which specializes in power plant and electric traction engineering. The interviewees hold the positions of a technical director and a head of custom software development. The following are the main issues they addressed:

- Modify questions regarding customers by differentiating between (external) customers and internal customers (i.e. employees), because they have two different business processes depending on type of customer;
- Modify questions regarding "would/do" by specifying what information is needed (does a company do something now or would the company do something in the future);
- Including "don't know" and "no" in the final survey.

3.2.5 The Czech Republic

To gain feedback on the pilot version of the survey form, the Prague team contacted Datlowe,s.r.o., which is one of the main commercial partners of the faculty in the field of Natural Language Processing. The interviewed person was an IT researcher experienced in data mining, but he possesses a good orientation in language processing too, so most questions were clear to him.

3.3 The final survey

3.3.1 Designing the final survey

The final survey was meant to be distributed electronically and to be self-explanatory so that respondents can complete it by themselves. The outcomes from the discussions about the observations from the pilot survey were integrated into the final survey. Originally Google Forms were considered for the distribution of the survey and the collection of answers. Unfortunately, Google Forms do not support multilingual survey. Instead, JotForms were used where multilinguality is supported. All answers are collected in one place irrespective of which language version the respondent uses which allows for all data to be used in the same analysis or report.

The language versions included the languages of the consortium partners: English, Slovene, Croatian, German and Czech.

The dissemination of the survey relied on the channels owned or used by consortium partners and personal contacts. Email lists and social media accounts (Facebook, LinkedIn, Twitter) proved only partially effective so personal contacts served as the main source of

responses. This made the procedure very time consuming and is also responsible for the number of responses which is below the anticipated goal of 200 responses.

3.3.2 Survey dissemination

3.3.2.1 Slovenia

With the conclusion of the pilot stage and subsequent modification of the questionnaire, the survey stage started in the middle of January 2017. The dissemination activities were being carried out until early March. For the Ljubljana part of the team, these were mainly in the form of queries via e-mail. 107 e-mail invitations to fill out the questionnaire were sent to a list of companies which was set up with the help of two online business registries: Slovenian Business Register (AJPES) and the list of members of the Technology Park Ljubljana.

In addition to these, approximately 30 companies were contacted through personal connections and asked to respond to our survey. The response rate from the latter group was very high, while the former's was very low. This explains the lower number of responses gathered than the set goal was. Nevertheless, we made efforts to diversify the responses in terms of what industries they represented and were fairly successful, managing to maintain a satisfactory level of relevance of our survey.

3.3.2.2 UK

Given that broader surveys such as the 2016 Expectations and Concerns of the European Language, an Industry Language Industry Survey with collaboration between Elia, European Master's in Translation, EUATC, GALA and LIND, received approx 400 responses, the number of results received from a more niche survey such as DigiLing is to be expected.
Leeds contributed to the dissemination of the Survey by posting it and encouraging responses on:

- the European Masters in Translation Facebook page, which has a community of 741 followers;
- the LeedsCTS Facebook Page, with a community of 1,693 followers;
- Elia Exchange representatives dissemination - The European Language Industry Association, Elia, is the premier trade association for the European language services industry embracing 200+ members throughout Europe and beyond. Elia Exchange is the leading programme from Elia that bridges the gap between the academic and business worlds for the benefit and prosperity of the language industry;
- Globalisation and Localization Association Forum (GALA Global), which has hundreds of members on five continents;
- Twitter accounts and hashtags: #LeedsCTS, #xl8, #t9n, #l10n, and @elearningbakery (1,392 followers)
- personal contacts and LinkedIn accounts of Leeds DigiLing team

3.3.2.3 Germany

The link was posted to social networks (LinkedIn, Twitter, Facebook) to elicit responses. The feedback expressed some criticism, for example one respondent inquired why the survey was not anonymous, another pointed out that they just needed more computational linguists. The posting resulted in 4 replies from Germany, 2 from Switzerland. We cannot say whether responses from other countries like the US, Syria or Vietnam, shortly after our post, resulted from the posting. On top of that, 7 people from the Language/Translation Technology community were approached via personal contacts, 4 of which had been briefed in person beforehand. None of them proceeded to fill out the survey, neither did they report that any of their contacts they had forwarded the survey to had expressed interest in the survey. Only one of them expressed interest in the topic/project.
Considering that Germany already has a well-developed Language and Translation Technologies industry, we assume that interest in the survey was bound to be low. This is, at the same time, a reminder for the German team to be take further action in advocating the importance of digital linguistics and differentiating it from computational linguistics, in particular from the point of view of future employers.

3.3.2.4 Croatia

Zagreb contributed to the dissemination of the Survey by posting it on the LinkedIn group of alumni of Faculty of Humanities and Social Sciences University of Zagreb (which has a community of 378 members), sending messages via LinkedIn to 60 personal contacts and emails to 6 personal contacts of Zagreb DigiLing Team. We encouraged our contacts to forward the survey to their contacts and to post it on mailing lists with potential survey respondents. We assume that most if not all feedback was from the personal contacts. If we take that assumption into consideration, the dissemination resulted in 23 respondents from Croatia and a response rate of 35%, an expected response rate for online surveys (Nulty 2008). We assume that the main reason for not participating in the survey is the lack of anonymity. The survey does not allow the respondents to leave out potentially sensitive data like name of the company and email. To be able to represent a company, one has to have clearance or hold higher positions. Acquiring clearance takes more time than filling in the survey, while for the second case we can presume that the higher the position in a company, the less time to fill in surveys.

3.3.2.5 The Czech Republic

First, the members of the team personally contacted (by phone or email) almost all commercial partners that cooperated with the Institute of Formal and Applied Linguistics in the last years. Second, we asked all other members of the Institute (which is some 60 people) to forward the survey to other potentially relevant organizations they are aware of. The response rate ranged from 100 % with the closest partners (with a commercial or academic cooperation in the last year), through some 30 % with we-had-once-a-project-in-common partners, to basically 0 % with all attempts without a prior personal contact. In total, the Prague team collected 20 completed responses from organizations ranging from

small publishing houses to large IT companies, of course most of them located in the Czech Republic.

3.4 Analysis of results

The survey of labour market needs was carried out in the months of January through March 2017. We received 81 responses from companies in eight different countries, with the large majority coming from the five countries of partner institutions. The size of the companies varied from very large (more than 5000 employees) to micro (1-19 employees), with fairly balanced distribution. With regards to industry sectors (Figure 5), roughly third (29) comes from the Technology sector, followed by Entertainment and Media (6), Communications (4), Healthcare (4) and Engineering and Construction (4). Other sectors were less well represented and 24 companies did not specify on this criteria. Some questions were mandatory while others where optional. All values in this chapter are given as percentages of actual answers received.

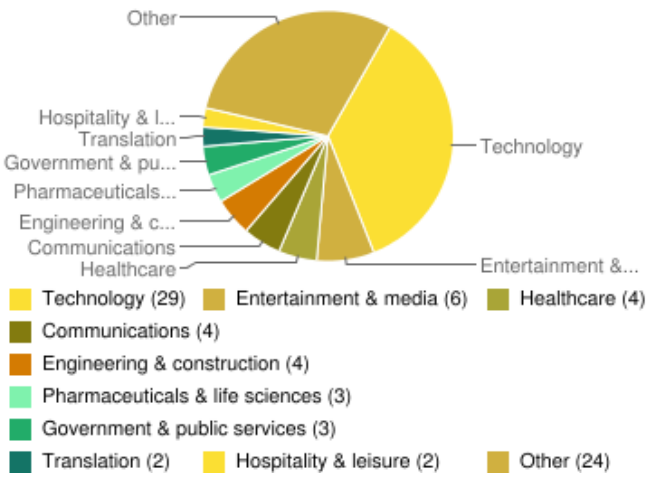


Figure 5: Respondents by industry

Looking at the results themselves, there is clear evidence of needs for automatic sorting, analysis and categorisation of e-mail communication and internal documents (68 and 73 percent of affirmative answers, respectively), while the need to automatically analyse social media or competitor's websites is less pronounced (Figure 6).

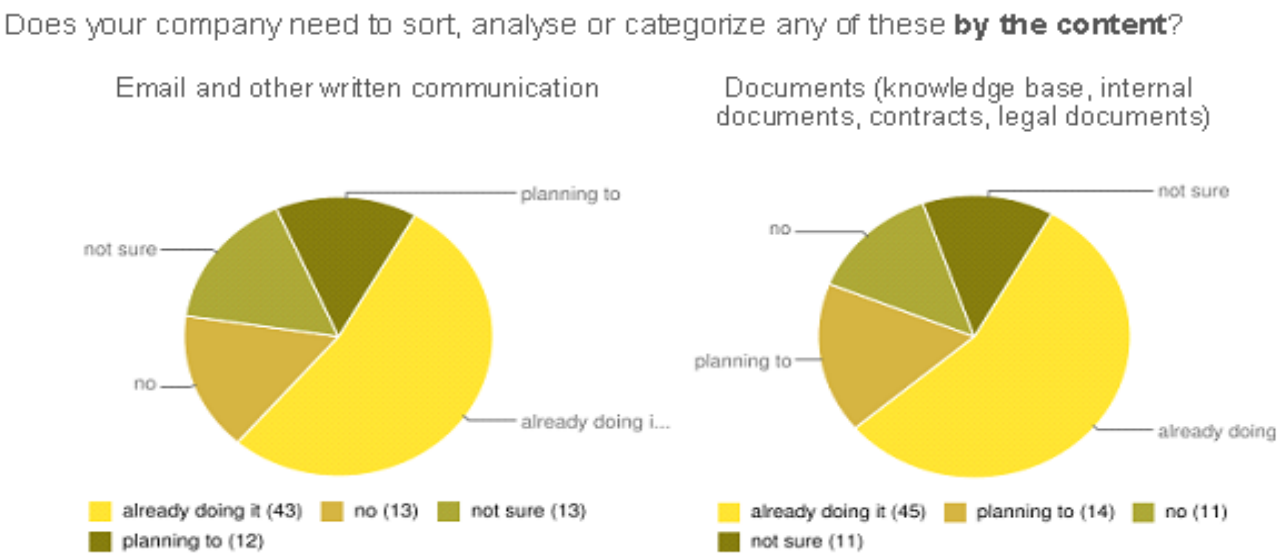


Figure 6: Needs for categorisation

Another trend that can not be overlooked is the need for multilinguality. In every relevant category (content, categorization, attitude analysis, terminology, automated communication and speech, digital agents), the need for automatic creation or processing of multilingual content was expressed by more than half the respondents, with terminology consistency checks figuring the highest at 75 percent (Figure 7).

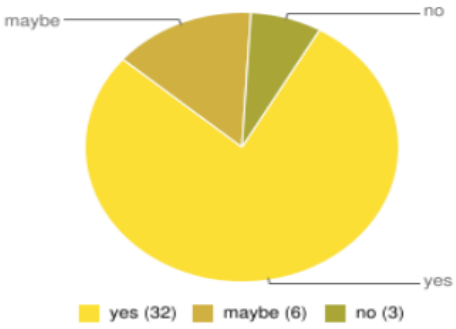


Figure 7: Need for multilingual terminology consistency checks

Clearly this represents significant opportunity, as well as challenge, for researchers and developers of terminology management solutions, machine translation technologies and even parallel corpora in general. Our survey's results seem to point in this direction and furthermore, it is our prediction that the need for multilinguality will continue to rise in the future, driven by trends of globalisation and internationalisation. Such predictions are again reaffirmed in the part of the survey concerning translation needs (Figure 8): 65 percent of interviewed companies need some form of translation now or will need it in future, with the most significant portion regularly translating to 1-6 foreign languages. Out of these, the largest portion still employ or outsource to human translators using CAT tools, while machine translation is used by roughly a quarter of respondents.

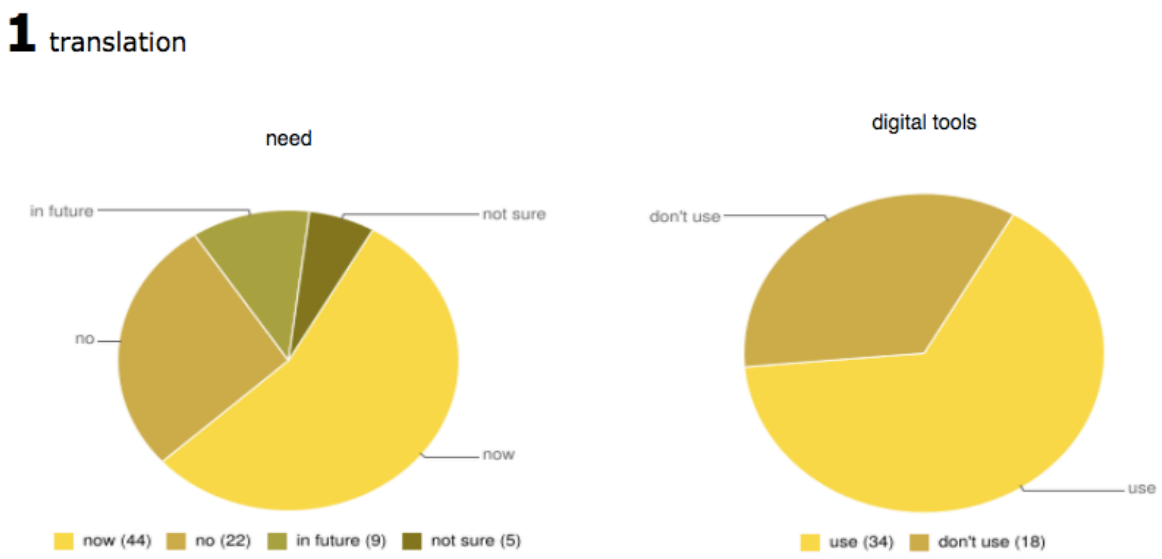


Figure 8: Need for translation and the use of digital tools

However, some of the answers also show that certain needs for automatic language technologies have still not taken root. For example, when asked about the need for automated responses to electronic communication (e-mails), only 36 percent answered affirmatively. Out of these, only half responded that they need it right now while the other half is merely considering it. The need for automated speech technologies for telephone communication with clients is even lower (16 percent). We may conclude from these responses that firstly, such technologies are still not developed enough to be trustworthy and, secondly, that communication with clients is deemed of highest importance by the interviewed companies and simply as something that can not yet be entrusted to machines.

4 Designing the model curriculum for Digital Linguistics

The DigiLing survey, despite the somewhat lower response rate than initially expected, reveals important trends regarding textual content processing and multilingual communication amongst European enterprises. The next task for the DigiLing consortium was to utilize these results to identify the key skills and competencies that a contemporary study programme at the academic level should provide in order for its graduates to be highly employable language professionals. A first step in this process is an analysis of existing Master's curricula at partner institutions which serves as the point of departure for the DigiLing model curriculum.

4.1 Existing Masters curricula at partner institutions

The DigiLing consortium comprises five academic partners from five EU countries. Each of them offers a number of Master's studies, but for the purposes of DigiLing we included only running Master's studies from the fields of translation, natural language processing and information/computer science. The total number of academic programmes included in this analysis is 12 (see Appendix 5).

Since Digital Linguistics is inherently an interdisciplinary field, partner institutions can be divided into two distinct groups by their focus. The first group is made up of Arts and Humanities institutes offering linguistic- and translation-oriented programmes – the Department for Translation Studies at the Faculty of Arts, University of Ljubljana, the Centre for Translation Studies at the University of Leeds and the Faculty of Arts at the Johannes Gutenberg University of Mainz all belong to this group. The second has a clear focus on technologies and language processing and is comprised of two partners: Department of Information and Communication Sciences at the Faculty of Humanities and Social Sciences, University of Zagreb, and the Faculty of Mathematics and Physics at the Charles University in Prague.

The divide between Humanities and Engineering is reflected also in the syllabi, still some topics or courses overlap and are part of both types of programmes. One such area is corpus studies, which can range from the completely applied approach (e.g. Corpus Linguistics for Translators offered by the Centre for Translation Studies, University in Leeds), to the more theoretical or specialised approach (e.g. Corpus Lexicography taught at the Department of Information and Communication Sciences, University of Zagreb).

Another common subject of study shared by most partner institutions is Machine Translation. Again, linguistics-oriented programmes take an applied and user-oriented approach to its study, such as in Principles and Applications of Machine Translation taught at the Centre for Translation Studies. On the other hand, IT-oriented programmes, e.g. the Computer Science – Computational Linguistics programme offered at the Faculty of Mathematics and Physics in Prague take a full-on theoretical approach to the study of MT with their course titled Statistical Machine Translation.

Continuing with this distinction – on the applied side of curricula we see many subjects related to the user aspect of language technologies. Examples of this can be found in many of the Centre for Translation Studies' modules, such as Computer-Assisted Translation (Applied Translation Studies programme), Computer and the Translator, Audiovisual Translation (Audiovisual Translation Studies programme), etc. The Faculty of Arts at the Johannes Gutenberg University in Mainz also offers such subjects as part of their Translation programme: Tools in Translation Process, Post Editing and Translation Project Management, to name a few. Finally, the Department of Translation Studies at the University in Ljubljana shows a similar picture: students have the chance to learn how to use language technologies in subjects such as Localisation, Subtitling and Translation Technologies. In addition to these, all of the partner institutions mentioned in this paragraph also offer traditional courses in both native and target language competence.

On the other end of the scale, students are taught not only how to use language tools, but how to actually *build* them. At the Department of Information and Communication Sciences,

University in Zagreb, they have the chance to learn segmentation and alignment for parallel corpora, automated syntax analysis, terminology extraction and similar. The Computer Science - Computational Linguistics programme at the Faculty of Mathematics and Physics of the Charles University in Prague offers subjects of similar depth. In Statistical Methods in NLP, for example, students learn generative and discriminative models, logistic regression, Bayesian networks, phrase-based and dependency-based statistical parsing, etc. In other subjects, they are being taught fundamentals of speech recognition and generation, supervised and unsupervised machine learning, Kernel functions, support vector machines and so on. There is also some general linguistics content, however the level of depth or the number of credits is – as might be expected – significantly lower than in Arts & Humanities programmes.

This overview leads us to the conclusion that there exists a gap between both kinds of studies which needs to be bridged if we want to train highly employable language professionals. The goal of our project is to implement a programme which would combine both sides and add the cognitive, psychological and sociological knowledge to create a new graduate profile, that of a Digital Linguist.

4.2 The DigiLing model curriculum

The overall learning outcome of the proposed curriculum is a highly skilled university graduate holding a Masters degree in Digital Linguistics who possesses knowledge and understanding about language and communication from several complementary disciplines. Understanding the inner workings of NLP systems enables our graduates to not only use them competently, but also detect potential flaws and suggest possible improvements. An understanding of the linguistic, pragmatic, intercultural, cognitive and sociological aspects of communication in the digital world enables them to create user-tailored content across various digital genres, communities and languages; coordinate teams which design and implement NLP systems; propose and help develop solutions for the processing or multilingual content; as well as continuously set the direction of further development by identifying and monitoring the needs of their users.

The Digital Linguistics curriculum is structured as follows (see Figure 8):

The entire curriculum is composed along three complementary pillars: Foundations, providing theoretical principles in each of the contributing fields: Linguistics, Multilingual Communication, Programing & IT and Digital Media. Since students will enter the programme with different Bachelors degrees and hence different backgrounds, the main purpose of this module is to equalize the differences and provide a solid interdisciplinary basis for all in accordance with their background.

The second pillar consists of Methods and Tools where students will acquire applied skills in selected areas of Digital Linguistics, including text analysis, digital content authoring, statistics, ethics & law and related. Some parts of this module will be obligatory, other

elective to broaden the scope of study. The main purpose of this module is to equip students with an inventory of methods and tools which will allow them to engage in research and applied projects developing own solutions to specific language-related problems.

The third pillar consists of applying the knowledge and skills to concrete problems, either through internship at partner companies/institutions or through projects developing the students' problem-solving and teamwork skills.

A Masters graduate in Digital Linguistics has the following skills and competencies:

- Language competence in at least two languages,
- An understanding of the way written and spoken language works at all levels of linguistic analysis,
- An understanding of the principles of multilingual communication, including skills in intercultural mediation, translation/interpreting/localization and multilingual content authoring,
- Skills in the compilation of digital language resources, such as corpora, lexica, acoustic databases and similar, including competencies in methodological design and technical implementation of LR compilation,
- Skills in analysing and processing natural language, including the ability to design and develop own tools as well as implement existing ones in order to analyse or process language data,
- Basic understanding of digital media from the sociological, psychological and legal perspective,
- Ability to perform independent research and acquire new skills,
- Ability to work in interdisciplinary/multilingual teams.

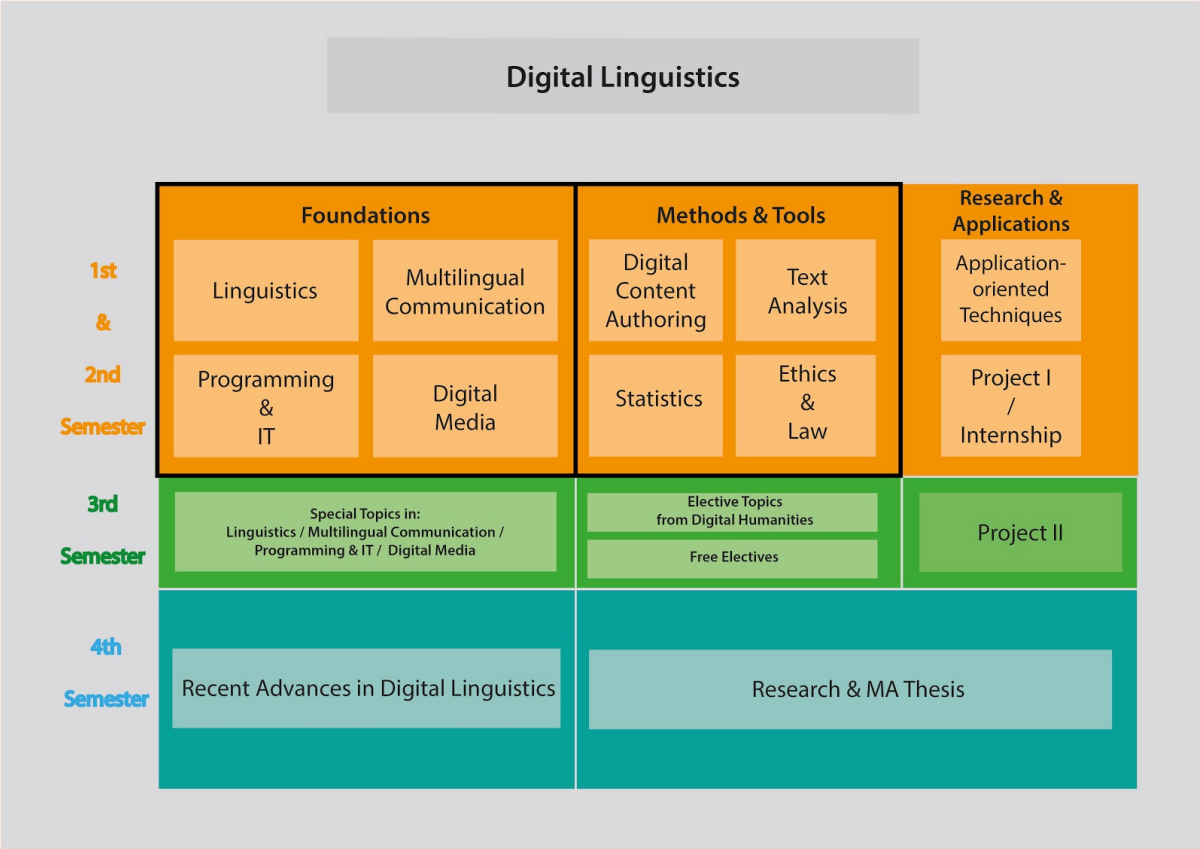


Figure 9: Structure of the curriculum in Digital Linguistics

As demonstrated in the previous sections, such an interdisciplinary array of educational and training activities cannot be provided by a single academic institution, even if co-operation between different faculties or institutes is established. Our model curriculum contains building blocks that as such do not yet exist at any of the partner universities, so they first need to be developed.

The DigiLing project will significantly contribute to the feasibility, sustainability and quality of a future Master's programme by producing a range of online modules covering many of the key topics in Digital Linguistics, including:

- Introduction to programming for linguists and students with a non-IT Bachelor's degree,
- Localisation workflows and project management,
- Basics of NLP for non-IT students,
- Machine Translation and post-editing,
- Basics of corpus linguistics with SketchEngine,
- Terminology management and mining.

The online modules will be accessible under the Creative Commons license through the DigiLing e-Learning platform and will be localized into all partner languages (except Czech) to facilitate inclusion into new national or joint study programmes. While the proposed model curriculum could be implemented at a single university, provided that the interdisciplinary areas are covered through the co-operation of different faculties or

institutes and that potential gaps could be filled through (partly supervised) e-learning modules, we believe that the best learning outcome can be achieved through an international collaboration network, preferably in the form of a joint degree.

4.3 Preparatory steps for a Joint Master's Degree in Digital Linguistics

While joint degrees represent the highest form of international co-operation between academic institutions, the process of preparing and consolidating such a programme within an international consortium is usually laborious and time-consuming, not to mention the process of accreditation which underlies national legislations and different – at times contradictory – regulations. Nevertheless, the DigiLing partners are committed to developing a Joint Master's degree in Digital Linguistics and proceed with all the steps necessary for its official accreditation and recognition.

The Joint Degree will involve four partner universities from Slovenia, Croatia, the Czech Republic and Germany. It is planned as a standard Master's programme lasting for two years and comprising 120 ECTS. International mobility of students is mandatory for the minimum of 1 semester, the equivalent of 30 ECTS. Apart from the students, we envisage strong co-operation and exchange between partners on the level of staff.

The development of the new programme comprises several stages:

- Consortium composition and official commitment to the programme (Letters of Intent)
- Syllabus design and consolidation
- Formal and legally binding specification of co-operation with all the administrative, financial, legal and programme-related issues (Partnership agreement)
- Accreditation by all partners

At the time of writing this report the first three of the above stages are already underway.

5 Conclusions

This deliverable marks the conclusion of the first stage of the DigiLing project, within which several important goals have been accomplished:

- **Existing surveys** and reports of the language industry and the trends within the market of language services have been **collected** and thoroughly **analysed**.

- **The DigiLing survey** aimed at enterprises outside of the LSP bubble has been **designed, localized and disseminated**. **81 responses were received** from eight countries. Despite the fact that we had been aiming for a larger number of responses, the number reached seems **representative** enough to draw relevant conclusions from the results.

- **Existing curricula at partner institutions** have been collected and put into a common database. A comparative analysis of 12 curricula was performed, which serves as one of the inputs for the **DigiLing model curriculum**.

- Gaps in the existing curricula together with the survey results helped DigiLing partners to **identify the target skills and competencies** of future Digital Linguists and to propose a list of modules to be developed as online courses for the **DigiLing e-Learning Hub.**

- A model curriculum for Digital Linguistics at the Master's level was designed and key preparatory steps for an **international interdisciplinary Joint Master's Degree** have been undertaken.

A publication of the results reported above in the form of a journal article is planned, as well as dissemination through the Erasmus+ project results platform and the DigiLing project website.

References

Gaspari et al., 2015. A survey of machine translation competences: Insights for translation technology educators and practitioners. Perspectives, 23:3, 333-358.

Language Industry Survey (LIS16). Expectations and Concerns of the European Language Industry. ELIA/EMT/EUATEC/GALA/LIND. Available at https://www.euatc.org/industry-surveys/item/403-2016-european-language-industry-survey-report.

Nulty, Duncan D, 2008. The adequacy of response rates to online and paper surveys: what can be done?. Assesment & Evaluation in Higher Education, 33:3, 301-314. https://www.uaf.edu/files/uafgov/fsadmin-nulty5-19-10.pdf

Pym, Anthony, 2013. Translation Skill-Sets in a Machine-Translation Age. Meta 583: 487–503.

SDL Language Technology Insights 2016 (SDL16). Available to registered users in the form of an executive summary at www.sdl.com.

Schmitt, Peter A., Lina Gerstmeyer, and Sarah Müller, 2016. Übersetzer und Dolmetscher: Eine internationale Umfrage zur Berufspraxis.

TAUS Translation Technology Report 2016 (TAUS16). Available to members of TAUS through www.taus.net.

Tsiavos, Prodromos, Stelios Piperidis, Maria Gavrilidou, Penny Labropoulou, Tasos Patrikakos, 2014. Language Resources - Legal Framework. Deliverable within the project Preparation and Launch of a Large-scale Action for Quality Translation Technology (QTLaunchPad).

Appendices

1. Pilot survey questionnaire
2. Final questionnaire (in English, Croatian, Czech, German, Slovene)
3. Summary of responses (as separate .xlsx file, anonymised)
4. Visual analysis of responses
5. Database of existing curricula at partner universities (as separate .xlsx file)

Appendix 1

Pilot survey

December 2016

Does your company need to process, analyse, categorize or search large quantities of any of these types of text data?

1. **emails + other written communication, internal and/or external**

   ☐ yes ☐ in future ☐ not sure, maybe not

2. **support tickets, customer queries**

   ☐ yes ☐ in future ☐ not sure, maybe not

3. **documents (knowledge base, internal documents, contracts, legal documents)**

   ☐ yes ☐ in future ☐ not sure, maybe not

4. **your own website(s) & social media**

   ☐ yes ☐ in future ☐ not sure, maybe not

5. **your competitor's website, social media, marketing emails/newsletters**

   ☐ now ☐ in future ☐ not sure, maybe not

List numbers in order of priority_____

Do you need people with expertise in "Google-like" technology for text searching, i.e. fast intelligent **searching** of large amounts of text able to include similar or related words and their different word forms.

need ☐ now ☐ in future ☐ not sure, maybe not

outsource*) ☐ definitely ☐ maybe or partly ☐ no, done in-house

multi-lingual**) ☐ yes ☐ maybe ☐ no

*) Do/Would you outsource or do internally? What proportion do you outsource? Would in-house experts be added value?
**) Would the technology need to accommodate more than one language?

What do you/would like to do with each type of data? What information/tasks do you/would like to retrieve/done done automatically?

1. automated **categorization** or **labelling** of documents, emails, support tickets or blog posts

   use ☐ now ☐ in future ☐ not sure, maybe not

   outsource*) ☐ definitely ☐ maybe or partly ☐ no, done in-house

   multi-lingual**) ☐ yes ☐ maybe ☐ no

2. automated analysis of **topics** and/or keywords in your website or competitor websites

   use ☐ now ☐ in future ☐ not sure, maybe not

   outsource*) ☐ definitely ☐ maybe or partly ☐ no, done in-house

   multi-lingual**) ☐ yes ☐ maybe ☐ no

3. automated analysis of **attitudes** in blog posts, comments and social media posts and media in general

   use ☐ now ☐ in future ☐ not sure, maybe not

   outsource*) ☐ definitely ☐ maybe or partly ☐ no, done in-house

   multi-lingual**) ☐ yes ☐ maybe ☐ no

4. consistency check of terminology or controlled language used in texts or databases

use	☐ now	☐ in future	☐ not sure, maybe
not outsource*)	☐ definitely	☐ maybe or partly	☐ no, done in-house
multi-lingual**)	☐ yes	☐ maybe	☐ no

List numbers in order of priority_____

Is there a need for **multimodal** / speech interfaces for the visually/hearing impaired?

- internal - employees
- external - clients/customers

*) Do/Would you outsource or do internally? What proportion do you outsource? Would in-house experts be added value?
**) Would the technology need to accommodate more than one language?

Do you need people with expertise in **automated communication** tasks?

1. automatic relevant responses to emails or other communications

use	☐ now	☐ in future	☐ not sure, maybe
not outsource*)	☐ definitely	☐ maybe or partly	☐ no, done in-house
multi-lingual**)	☐ yes	☐ maybe	☐ no

2. **chat bots** for automated website chat communication

use	☐ now	☐ in future	☐ not sure, maybe
not outsource*)	☐ definitely	☐ maybe or partly	☐ no, done in-house
multi-lingual**)	☐ yes	☐ maybe	☐ no

3. automated **speech / dialogue** for phone communications with clients

use	☐ now	☐ in future	☐ not sure, maybe

not
outsource*) ☐ definitely ☐ maybe or partly ☐ no, done in-house
multi-lingual**) ☐ yes ☐ maybe ☐ no

4. animated digital agents (avatars) on websites
   use ☐ now ☐ in future ☐ not sure, maybe
   not
   outsource*) ☐ definitely ☐ maybe or partly ☐ no, done in-house
   multi-lingual**) ☐ yes ☐ maybe ☐ no

List numbers in order of priority_____

*) Do/Would you outsource or do internally? What proportion do you outsource? Would in-house experts be added value?
**) Would the technology need to accommodate more than one language?

Does/Would your company need/already use any of the below? Would you be interested in having experts able to exploit digital tools to help speed up the process and improve quality and consistency of…?

1. human translation
   need ☐ now ☐ in future ☐ not sure, maybe
   not
   digital tools ☐ use ☐ don't use
   outsource*) ☐ yes ☐ no

2. machine translation
   need ☐ now ☐ in future ☐ not sure, maybe
   not
   outsource*) ☐ yes ☐ maybe or partly ☐ no, done in-house

3. web content authoring
   need ☐ now ☐ in future ☐ not sure, maybe
   not

outsource*)	☐ definitely	☐ maybe or partly	☐ no, done in-house
multi-lingual**)	☐ yes	☐ maybe	☐ no

4. copywriting

need not	☐ now	☐ in future	☐ not sure, maybe
outsource*)	☐ definitely	☐ maybe or partly	☐ no, done in-house
multi-lingual**)	☐ yes	☐ maybe	☐ no

5. writing assistants

need not	☐ now	☐ in future	☐ not sure, maybe
outsource*)	☐ definitely	☐ maybe or partly	☐ no, done in-house
multi-lingual**)	☐ yes	☐ maybe	☐ no

6. brand name research & development

need not	☐ now	☐ in future	☐ not sure, maybe
outsource*)	☐ definitely	☐ maybe or partly	☐ no, done in-house
multi-lingual**)	☐ yes	☐ maybe	☐ no

List numbers in order of priority_____

Interviewer's notes

Target entity types

localization companies
chamber of commerce

recruiting agencies
companies:
inbound marketing companies
digital marketing + SEO
media
information extraction
legal companies + legal depts of companies
document classification (medical, legal)

Aim

The aim of this pilot questionnaire is:

- to find out if the wording is understandable
- whether the questions guide the respondents to useful answers

How to

Ask the questions and if the respondent is confused, provide the hints in Digiling green, focusing on the **benefits** to the company. Avoid using jargon (lemmatized, keyword extraction, sentiment analysis etc.) Imagine talking to your best non-NLP friend (if you have any :-)

Hints to elicit better answers

page1

Does your company need to analyse, categorize or search large quantities of any of these types of text data?

6. emails + other written communication, internal and/or external
   apparently all companies will have communication of this type but ask about the quantities and whether they need to search or analyse or extract data from a large archive, eg. 100,000s of emails
7. support tickets
   most will have this too but again, ask about searching, analysing large archives

8. documents (internal documents, contracts, legal documents)
   not all of them may have large document repositories, ask if they do
9. your own website(s) & social media
   explain that understanding the content of your own website or external communication can be exploited for marketing purposes
10. competitor's websites & social media, marketing emails/newsletters
    explain that automatic analysis of the language used can shed light on the marketing strategy, how they refer to their products, what language they use, the register etc.

Do you need people with expertise in "Google-like" technology for text searching, i.e. fast intelligent searching of large amounts of text able to include similar or related words and their different word forms.
explain that search performance can be greatly improved by employing similar technology to what they know from google and this can be implemented locally inside their corporate system to preserve confidentiality

page2

What do you/would like to do with each type of data? What information/tasks do you/would like to retrieve/done done automatically?

5. automated **categorization** or **labelling** of documents, emails, support tickets or blog posts
   tagging and categorizing can be done automatically to allow for systematic processing or analysis, e.g. customer service evaluation
   emails can be automatically assigned to correct people by a system which uses machine learning to improve performance with time
6. automated analysis of **topics** and/or keywords in your website or competitor websites
   words specific to the text can be automatically extracted to group texts with similar content or to categorize or label as above
   a system can be set up to monitor competitor's website or online materials and detect any changes in the way the product is presented
7. automated analysis of **attitudes** in blog posts, comments and social media posts and media in general
   texts can be automatically analysed to detect positive or negative attitudes of the author about your own or competitor's product leading to an automatic in-depth analysis of customer's opinions
8. consistency check of terminology or controlled language used in texts or databases
   terminology used in manuals, internal or external documents can be checked for consistency, terms or formulations not approved by the company can be automatically detected

Is there a need for **multimodal** / speech interfaces for the visually/hearing impaired?

- internal - employees
- external - clients/customers
  e.g. tablets given to the hearing impaired clients at a bank with speech-to-text technology

page3

Do you need people with expertise in **automated communication** tasks?

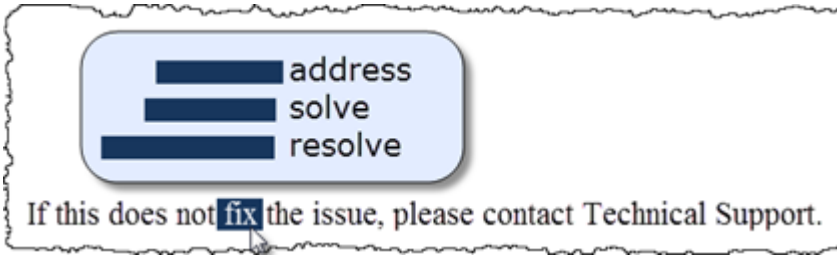
5. automatic relevant responses to emails or other communications
   explain that such an email can automatically select the best fitting link to the knowledge base or even contain a personalized, automatically generated reply
6. **chat bots** for automated website chat communication
   as many businesses now opt for a chat pop-up as a means of communication, automated bots could be used for the initial exchange of basic information before being handed to a live interlocutor or could even handle the complete communication dependent on the complexity of the query
7. automated **speech / dialogue** for phone communications with clients
   as with chat bots, explain that the system can be designed to give personalized replies, not just pre-recorded ones
8. animated digital agents (avatars) on websites
   as with chat bots

page4

Does/Would your company need/already use any of the below? Would you be interested in having experts able to exploit digital tools to help speed up the process and improve quality and consistency?

7. human translation
   explain that human translation can be aided by various digital tools helping the translator check the usage, check terminology, look up translation suggestions from previously translated material, auto-translate stretches of text translated in the past etc.
8. machine translation
   explain that the quality of machine translation can be hugely aided by exploiting corporate document repository for which an in-house system has to be designed to preserve confidentiality
9. web content authoring
   checking the density of keywords, term consistency, suggestions for creative language (distributional thesaurus)

10. copywriting
    getting suggestions from competitor texts or getting suggestions from a distributional thesaurus for using words creatively
11. writing assistants
    writing suggestions (see img) can significantly improve the quality of language, both mother tongue and foreign language



12. brand name research & development
    suggestions for related words including distantly related are invaluable of brand name research

Appendix 2: Final survey

- English
- Croatian
- Czech
- German
- Slovene

DIGI / LING

TRANS / EUROPEAN / E-LEARNING

HUB / FOR / DIGITAL / LINGUISTICS

Help us educate **digital experts** for you

About DigiLing

www.digiling.eu

DigiLing is an Erasmus+ project with a goal of meeting the increasing need for experts in processing, analyzing, searching and exploiting large amounts of text such as email communication, support tickets, internal documentation, websites, social media or media in general. DigiLing will produce e-learning materials covering the essential skills and competencies in Digital Linguistics to help educate **Digital Linguists** for companies such as yours.

A **Digital Linguist** is a specialist who can help your company become more visible, engage more effectively with your user communities, understand your local and international market trends and competition better, and generally make even more of your existing text data.

The following survey aims to find out the extent to which you already engage in the activities mentioned above, or the degree to which you need a Digital Linguist as part of your team. It will also enable us to shape the training of Digital Linguists, and create learning materials aimed to develop those specific skills the industry values.

Company information

company name _____

company size
1 - 19 employees
20 -99 employees
100-499 employees
500-999 employees
1,000-4,999 employees
more than 5,000 employees

corporate website _____

Country _____ city _____

respondent name _____ email _____

Industry

☐ Aerospace, defence & security
☐ Automotive
☐ Banking & capital markets
☐ Chemicals
☐ Communications
☐ Energy, utilities & mining
☐ Engineering & construction
☐ Entertainment & media
☐ Financial services
☐ Forest, paper & packaging
☐ Government & public services
☐ Healthcare
☐ Hospitality & leisure
☐ Industrial manufacturing
☐ Insurance
☐ Pharmaceuticals & life sciences
☐ Retail & consumer
☐ Technology
☐ Transportation & logistics
☐ Other (please specify)_____

How can a Digital Linguist help your company grow?

Does your company need to sort, analyse or categorize any of these by the **content**?

- **emails** and other **written communication**
  ☐already doing it ☐planning to ☐not sure ☐no

- support **tickets**, customer **queries**
  ☐already doing it ☐planning to ☐not sure ☐no

- **documents** (knowledge base, internal documents, contracts, legal documents)
  ☐already doing it ☐planning to ☐not sure ☐no

- your own or competitor's **website**(s), **social media** or **marketing** emails
  ☐already doing it ☐planning to ☐not sure ☐no

Is there a need for sophisticated text searching, i.e. fast intelligent **searching** of large amounts of text able to include similar or related words and their different word forms.

need	☐now	☐in future	☐not sure	☐no
outsource*)	☐definitely	☐maybe	☐no, done in-house	
multi-lingual**)	☐yes	☐maybe	☐no	

What do you do with the data? Who performs these tasks?

1. **categorization** by topics, **labelling** or **tagging**

when?	☐doing it already	☐in future	☐not sure	☐no
outsource*)	☐definitely	☐maybe or partly	☐no, done in-house	
multi-lingual**)	☐yes	☐maybe	☐no	

2. analysis of **attitudes** in blog posts, comments and social media posts and media in general

when?	☐doing it already	☐in future	☐not sure	☐no
outsource*)	☐definitely	☐maybe or partly	☐no, done in-house	
multi-lingual**)	☐yes	☐maybe	☐no	

3. consistency check of **terminology** or controlled **language** used in texts or databases

when?	☐doing it already	☐in future	☐not sure	☐no
outsource*)	☐definitely	☐maybe or partly	☐no, done in-house	
multi-lingual**)	☐yes	☐maybe	☐no	

List numbers in order of priority (comma separated numbers) _____

Any comments on the questions on this page? _____

*) Would you design your own system and/or use your own specialists or would you get a ready-made solution?

**) Would the technology need to accommodate more than one language?

Do you need people with expertise in **automated communication** technologies such as:

1. automatic relevant **responses** to emails or other communication (possibly before handing over to a live agent)

need	☐now	☐in future	☐not sure, no
outsource*)	☐definitely	☐maybe	☐no, done in-house
multi-lingual**)	☐yes	☐maybe	☐no

2. **chat bots** for automated website chat communication (possibly to only get some basic information before handing over to a live agent)

need	☐now	☐in future	☐not sure, no
outsource*)	☐definitely	☐maybe or partly	☐no, done in-house
multi-lingual**)	☐yes	☐maybe	☐no

3. automated **speech / dialogue** for phone communications with clients

need	☐now	☐in future	☐not sure, no
outsource*)	☐definitely	☐maybe or partly	☐no, done in-house
multi-lingual**)	☐yes	☐maybe	☐no

4. animated digital agents (avatars) on websites

need	☐now	☐in future	☐not sure, no
outsource*)	☐definitely	☐maybe or partly	☐no, done in-house
multi-lingual**)	☐yes	☐maybe	☐no

List numbers in order of priority (comma separated numbers) _____

How do your visually/hearing impaired **employees** use your systems and interact with your web presence?

_____

How do your visually/hearing impaired **clients** interact with your web presence?

_____

Any comments on the questions on this page? _____

*) Would you design your own system and/or use your own specialists or would you get a ready-made solution?
**) Would the technology need to accommodate more than one language?

Does your company regularly need people and/or tools for…?

1. translation

need	☐now	☐in future	☐not sure, maybe not
digital tools	☐use	☐don't use	
outsource*)	☐yes	☐no	

[if need:now]
How many languages do you regularly translate into?
Estimated volume of pages per week?
[if digital tools:use]
Which tools do you use:
translation productivity tools (CAT tools)
machine translation
other (please specify):_____
[if outsource:no]
How many persons does your company employ for translation tasks?

2. web content authoring

need	☐now	☐in future	☐not sure, maybe not
outsource*)	☐definitely	☐maybe or partly	☐no, done in-house
multi-lingual**)	☐yes	☐maybe	☐no

3. copywriting

need	☐now	☐in future	☐not sure, maybe not
outsource*)	☐definitely	☐maybe or partly	☐no, done in-house
multi-lingual**)	☐yes	☐maybe	☐no

List numbers in order of priority _____

Is there a use case for **writing assistants** - a tool providing feedback as the user types and offering suggestions for better wording or correct use of terminology based on the company standards

need	☐now	☐in future	☐not sure, maybe not
outsource*)	☐definitely	☐maybe or partly	☐no, done in-house
multi-lingual**)	☐yes	☐maybe	☐no

Any comments on the questions on this page? _____

*) Would you design your own system and/or use your own specialists or would you get a ready-made solution?
**) Would the technology need to accommodate more than one language?

0%
0% 0 / 0 Required Fields Complete.

Czech

Pomozte nám vychovat lepší IT odborníky pro vás

O projektu DigiLing

www.digiling.eu

DigiLing je projekt programu Erasmus+, který reaguje na stále se zvyšující zájem trhu práce o odborníky na zpracování, analýzu a prohledávání velkého objemu textových dat, jako jsou například emailová komunikace, databáze zákaznické podpory, webový obsah, sociální sítě nebo média obecně. Projekt DigiLing má za cíl připravit e-learningové výukové materiály zaměřené na počítačovou lingvistiku, které budou využity pro přípravu **počítačových lingvistů**, kteří najdou využití ve firmách, jako je ta vaše.

Počítačový lingvista je odborník, který pomáhá firmě lépe se zviditelnit, efektivněji komunikovat se zákazníky, lépe porozumět změnám na trhu a, obecně řečeno, maximálně využít existující data.

Cílem tohoto průzkumu je zjistit, do jaké míry vaše firma již využívá výše uvedené možnosti a zda by počítačový lingvista mohl vaší firmě pomoci. Tyto informace nám pomohou vytvořit výukové materiály, které vybaví odborníky dovednostmi ceněnými právě ve vašem odvětví.

Partneři projektu DigiLing

University of Ljubljana

JG|U JOHANNES GUTENBERG UNIVERSITÄT MAINZ

UNIVERSITAS STUDIORUM ZAGRABIENSIS MDCLXIX

Informace o vaší firmě

název firmy *

webová stránka *

velikost firmy *

- 1 - 19 zaměstnanců
- 20 - 99 zaměstnanců
- 100 - 499 zaměstnanců
- 500 - 999 zaměstnanců
- 1 000 - 4 999 zaměstnanců
- více než 5 000 zaměstnanců

země *

- Slovinsko
- Chorvatsko
- Velká Británie
- Německo
- Česká Republika
- jiná

město *

respondent

křestní jméno | příjmení

E-mail *

Odvětví *

- Aerospace, defence & security
- Automotive
- Banking & capital markets
- Chemicals
- Communications
- Energy, utilities & mining
- Engineering & construction
- Entertainment & media
- Financial services
- Forest, paper & packaging
- Government & public services
- Healthcare
- Hospitality & leisure
- Industrial manufacturing
- Insurance
- Pharmaceuticals & life sciences
- Retail & consumer
- Technology
- Transportation & logistics
- jiná

Potřebuje vaše firma třídit nebo analyzovat následující typy dat **dle obsahu**?

emaily a další písemná komunikace *

- již provádíme
- plánujeme
- nevím
- ne

případy (tickety) v systému zákaznické podpory *

- již provádíme
- plánujeme
- nevím
- ne

dokumenty (znalostní báze, interní dokumentace, smlouvy, právní dokumenty) *

- již provádíme
- plánujeme
- nevím
- ne

vlastní nebo konkurenčí webové stránky, sociální sítě nebo emaily s obchodním sdělením *

Má vaše firma využití pro systém pokročilého prohledávání textu, tj. rychlé inteligentní prohledávání velkých objemů dat s možností automaticky zahrnout synonyma nebo různé slovní tvary téhož slova?

potřebujeme *

- nyní
- v budoucnosti
- nevím
- ne

Jak s těmito daty nakládáte? Kdo tyto úkoly provádí?

1 **třídění** do kategorií nebo přiřazování značek **dle obsahu**

kdy? *

- již provádíme
- v budoucnosti
- nevím
- ne

2 analýza **postojů** v blogových příspěvcích, komentářích, sociálních sítích a médiích obecně

kdy? *

- již provádíme
- v budoucnosti
- nevím
- ne

Jak s těmito daty nakládáte? Kdo tyto úkoly provádí?

3 kontrola jednotnosti **terminologie** a standardizovaného jazyka ve firemních textech a databázích

kdy? *

Seřaďte prosím tyto body dle důležitosti.

1 třídění do **kategorií** nebo přiřazování **značek** dle obsahu

2 analýza postojů v blogových příspěvcích, komentářích, sociálních sítích a médiích obecně

3 kontrola jednotnosti terminologie a standardizovaného jazyka ve firemních textech a databázích

pořadí důležitosti * např. 312

od nejdůležitějšího po nejméně důležité

Sem prosím vepište jakékoli doplňující informace k výše uvedeným tématům.

Potřebujete odborníky se specializací na systémy **automatizované komunikace**?

1 systém pro automatické relevantní **odpovědi** na emaily či jinou komunikaci (například před předáním komunikace živému agentovi)

potřebujeme *

- nyní
- v budoucnosti
- nevím
- ne

2 **Chatboti** pro automatickou komunikaci na chatu firemních webových stránek (například za účelem zjištění základních informací před předáním konverzace živému agentovi)

Potřebujete odborníky se specializací na systémy **automatizované komunikace**?

3 automatizovaný telefonický dialog se zákazníky

potřebujeme *

- ○ nyní
- ○ v budoucnosti
- ○ nevím
- ○ ne

4 **animovaní asistenti** (avatar) na webových stránkách

potřebujeme *

- ○ nyní
- ○ v budoucnosti
- ○ nevím
- ○ ne

Seřaďte prosím dle důležitosti.

1 systém pro automatické relevantní **odpovědi** na emaily či jinou komunikaci

2 Chatboti pro automatickou komunikaci na chatu firemních webových stránek

3 automatizovaný **telefonický dialog** se zákazníky

4 animovaní asistenti (avatar) na webových stránkách

pořadí důležitosti * např. 312

od nejdůležitějšího po nejméně důležité

Sem prosím vepište jakékoli doplňující informace k výše uvedeným tématům.

Jak vaši zrakově nebo sluchově postižení klienti/zákazníci pracují s vašimi webovými stránkami? *

Jakékoli poznámky k výše uvedeným tématům vepiště prosím sem.

Potřebujete pravidelně využívat následující odborníky nebo nástroje na…?

1 jazykové překlady

potřebujeme *

- ○ nyní
- ○ v budoucnosti
- ○ nevím
- ○ ne

2 tvorba obsahu webových stránek

potřebujeme *

- ○ nyní
- ○ v budoucnosti
- ○ nevím
- ○ ne

potřebujeme *

- nyní
- v budoucnosti
- nevím
- ne

Uveďte v pořadí důležitosti

1 jazykové překladay

2 tvorba obsahu webových stránek

3 reklamní a marketingové texty

pořadí důležitosti * např. 312
od nejdůležitějšího

Sem prosím vepište jakékoli doplňující informace k výše uvedeným tématům.

Najde ve vaší firmě využití **asistent pro psaní** - nástroj, který poskytuje zpětnou vazbu při psaní textu, nabízí volby vhodnějších slov a použití správné terminologie dle firemních standardů.

potřebujeme *

- nyní
- v budoucnosti
- nevím
- ne

Jakékoli doplňující informace a poznámky vepište prosím sem.

Pomozite nam obrazovati **digitalne stručnjake**

O projektu DigiLing

www.digiling.eu

DigiLing je Erasmus+ projekt čiji cilj je zadovoljiti sve veću potrebu za stručnjacima u obradi, analizi, pretraživanju i upotrebi velikih količina tekstualnih podataka, kao što su e-pošta, podrška za korisnike i/li kupce, interna dokumentacija, mrežne stranice, društvene mreže ili mediji općenito. Cilj je projekta stvoriti materijale za e-učenje koji pokrivaju osnovne vještine i znanja iz digitalne lingvistike, usmjerene obrazovanju **digitalnih lingvista** za poduzeća kao što je Vaše.

Digitalni lingvist stručnjak je koji može pomoći Vašem poduzeću poboljšati prepoznatljivost poduzeća, učinkovitije sudjelovanje u zajednicama korisnika i kupaca, bolje razumijevanje konkurenata te lokalnih i međunarodnih trendova tržišta. Općenito, digitalni lingvist pomoći će vam bolje iskoristiti vaše postojeće tekstualne podatke.

Sljedećom anketom želimo ustanoviti provodite li već sada bar djelomično neke od gore navedenih aktivnosti te koliko će digitalni lingvisti doprinijeti vašem poslovanju. Također će nam omogućiti usmjeriti obrazovanje digitalnih lingvista te stvoriti materijale za učenje koji će razvijati specifične vještine koje gospodarstvo najviše vrednuje.

Podaci o poduzeću

Naziv poduzeća _____

Veličina poduzeća (broj zaposlenika)
1 - 19
20 -99
100-499
500-999
1.000-4.999
više od 5.000

Mrežna stranica poduzeća _____

Država _____ Grad _____

Ime ispitanika (neobavezno) _____ E-pošta _____

Industrija

☐ Zrakoplovstvo, obrana i sigurnost
☐ Automobilska industrija
☐ Bankarstvo i tržište kapitala
☐ Kemijska industrija
☐ Komunikacije
☐ Energetska industrija, rudarstvo i komunalna postrojenja
☐ Strojarska i građevinska industrija
☐ Zabava i mediji
☐ Financijske usluge
☐ Šumarstvo, papir i ambalaža
☐ Državne i javne službe
☐ Zdravstvo
☐ Turizam i slobodno vrijeme
☐ Industrijska proizvodnja
☐ Osiguranje
☐ Farmaceutska industrija i biološke znanosti
☐ Maloprodaja i trgovina
☐ Tehnologija
☐ Transportna industrija i logistika
☐ Ostalo (molimo navesti) _____

Kako može digitalni lingvist pomoći vašem poduzeću?

Treba li vašem poduzeću razvrstavanje, analiza ili kategorizacija nekog od ovih **sadržaja**?

- **e-pošta** i ostala **pisana komunikacija**
  ☐ već radimo ☐ planiramo raditi ☐ nismo sigurni ☐ ne

- **podrška** korisnicima ili kupcima
  ☐ već radimo ☐ planiramo raditi ☐ nismo sigurni ☐ ne

- **dokumenti** (baze znanja, interni dokumenti, ugovori, pravni dokumenti)
  ☐ već radimo ☐ planiramo raditi ☐ nismo sigurni ☐ ne

- **mrežne stranice** vlastitog ili konkurentskog poduzeća, **društveni mediji** ili **marketinška** e-pošta
  ☐ već radimo ☐ planiramo raditi ☐ nismo sigurni ☐ ne

Postoji li potreba za naprednim pretraživanjem teksta, odnosno brzim i pametnim **pretraživanjem** velikih količina tekstova koje je sposobno obuhvatiti slične ili srodne riječi te njihove različite oblike

trebamo	☐ sada	☐u budućnosti	☐nismo sigurni	☐ne
*outsourcing**)	☐obavezno	☐možda	☐ne, interno	
višejezičnost**)	☐da	☐možda	☐ne	

Što radite s podacima? Tko obavlja te zadatke?

1. **kategorizacija** po temama ili **označavanje** dokumenata

kada?	☐već radimo	☐u budućnosti	☐nismo sigurni	☐ne
*outsourcing**)	☐obavezno	☐možda ili djelomično	☐ne, interno	
višejezičnost**)	☐da	☐možda	☐ne	

2. analiza **stavova** u blogovima, komentarima i društvenim mrežama te medijima općenito

kada?	☐već radimo	☐u budućnosti	☐nisam siguran	☐ne
*outsourcing**)	☐obavezno	☐možda ili djelomično	☐ne, interno	
višejezičnost**)	☐da	☐možda	☐ne	

3. provjera **dosljednosti termina** ili upotreba **kontroliranog jezika** u tekstovima ili bazama podataka

kada?	☐već radimo	☐u budućnosti	☐nismo sigurni	☐ne
*outsourcing**)	☐obavezno	☐možda ili djelomično	☐ne, interno	
višejezičnost**)	☐da	☐možda	☐ne	

Molimo poredajte brojeve prema važnosti (razdvojite ih zarezom) _____

Imate li komentare na ovu stranicu? _____

*) Biste li razvili vlastiti sustav i/li angažirali vlastite stručnjake ili biste kupili gotova rješenja?

**) Treba li tehnologija podržavati više od jednog jezika?

Trebaju li vam stručnjaci koji razvijaju tehnologiju **automatske komunikacije**, kao što su:

1. automatsko pametno **odgovaranje** na e-poštu i drugu komunikaciju (u potpunosti ili kao korak prije predaje živom agentu)

trebamo	☐sada	☐u budućnosti	☐nismo sigurni, ne
*outsourcing**)	☐obavezno	☐možda	☐ne, interno
višejezičnost**)	☐da	☐možda	☐ne

2. **robotski sugovornik** za automatsku komunikaciju na mrežnim stranicama (u potpunosti ili kao korak prije predaje živom agentu)

trebamo	☐sada	☐u budućnosti	☐nismo sigurni, ne
*outsourcing**)	☐obavezno	☐možda ili djelomično	☐ne, interno
višejezičnost**)	☐da	☐možda	☐ne

3. sintetiziran **govor/dijalog** za telefonsku komunikaciju s kupcima ili korisnicima

trebamo	☐sada	☐u budućnosti	☐nismo sigurni, ne
*outsourcing**)	☐obavezno	☐možda ili djelomično	☐ne, interno
višejezičnost**)	☐da	☐možda	☐ne

4. **animirani agenti** (avatari) na mrežnim stranicama

trebamo	☐sada	☐u budućnosti	☐nismo sigurni, ne
*outsourcing**)	☐obavezno	☐možda ili djelomično	☐ne, interno
višejezičnost**)	☐da	☐možda	☐ne

Molimo poredajte brojeve prema važnosti (razdvojite ih zarezom) _____

Kako vaši **zaposlenici** oštećena vida/sluha koriste vaše sustave i mrežne stranice?

_____

Kako vaši **kupci i/li korisnici** oštećena vida/sluha koriste vaše mrežne stranice?

_____

Imate li komentare na ovu stranicu? _____

*) Biste li razvili vlastiti sustav i/li angažirali vlastite stručnjake ili biste kupili gotova rješenja?
**) Treba li tehnologija podržavati više od jednog jezika?

Trebaju li vašem poduzeću redovito ljudi i/li alati za…?

1. prevođenje
   trebamo ☐sada ☐u budućnosti ☐nismo sigurni, možda ne
   digitalni alati ☐upotrebljavamo ☐ne upotrebljavamo
   *outsourcing**) ☐da ☐ne
   [ako ste odgovorili da "sada" trebate ljude i/li alate za]
   Na koliko jezika najčešće trebate prijevode?
   Približno koliko stranica na tjedan?
   [ako ste odgovorili da "upotrebljavate" digitalne alate]
   Koje alate upotrebljavate:
   računalno potpomognuto prevođenje (CAT alati)
   strojno prevođenje
   ostalo (molimo navedite):_____
   [ako ste odgovorili da "ne" trebate outsourcing]
   Koliko osoba zapošljava vaše poduzeće za prevođenje?

2. stvaranje mrežnog sadržaja
   trebamo ☐sada ☐u budućnosti ☐nismo sigurni, možda ne
   *outsourcing**) ☐obavezno ☐možda ili djelomično ☐ne, interno
   višejezičnost**) ☐da ☐možda ☐ne

3. stvaranje oglasnih i promotivnih materijala
   trebamo ☐sada ☐u budućnosti ☐nismo sigurni, možda ne
   *outsourcing**) ☐obavezno ☐možda ili djelomično ☐ne, interno
   višejezičnost**) ☐da ☐možda ☐ne

Molimo poredajte brojeve prema važnosti (razdvojite ih zarezom) _____

Trebate li **digitalnog pomoćnika za pisanje** (*writing assistants*) - alat koji prilikom pisanja redovito nudi prijedloge za bolje izražavanje i nadzire pravilno korištenje terminologije u skladu sa standardima vašeg poduzeća?
trebamo ☐sada ☐u budućnosti ☐nismo sigurni, možda ne
*outsourcing**) ☐obavezno ☐možda ili djelomično ☐ne, interno
višejezičnost**) ☐da ☐možda ☐ne

Imate li komentare na ovu stranicu? _____

*) Biste li razvili vlastiti sustav i/li angažirali vlastite stručnjake ili biste kupili gotova rješenja?
**) Treba li tehnologija podržavati više od jednog jezika?

DIGI / LING

TRANS / EUROPEAN / E-LEARNING

HUB / FOR / DIGITAL / LINGUISTICS

Helfen Sie uns dabei, **Experten für Ihre digitale Kommunikation** auszubilden

Über DigiLing

www.digiling.eu

DigiLing ist ein Projekt des Programms Erasmus+ mit dem Ziel, den steigenden Bedarf an Fachkräften zu decken, die große Textmengen verarbeiten, analysieren, durchsuchen und erschließen können, wie sie z. B. bei E-Mail-Kommunikation, Supportanfragen, interner Dokumentation, auf Webseiten, in soziale Medien oder anderen Medien auftreten können. DigiLing hat sich zum Ziel gesetzt, E-Learning-Materialien für den Erwerb grundlegender Fähigkeiten und Kompetenzen der digitalen Linguistik zu erstellen, um die **Experten der digitalen Linguistik** von morgen auszubilden.

Experten der digitalen Linguistik können Ihrem Unternehmen helfen, in der digitalen Welt sichtbarer zu werden, effektiver mit Ihren Nutzergemeinschaften zu interagieren, lokale und internationale Trends und Konkurrenzen auf den Märkten besser zu verstehen, und allgemein noch mehr aus Ihren vorhandenen Text-Daten herauszuholen.

Die folgende Umfrage soll festhalten, inwiefern Sie bereits in den oben genannten Feldern tätig sind und wie sehr Sie einen Experten der digitalen Linguistik in Ihrem Team brauchen. Außerdem wird die Umfrage uns helfen, die Ausbildung von digitalen Linguisten voranzubringen und Lernmaterialien zu gestalten, die die Kenntnisse vermitteln, die in der Industrie gebraucht werden.

Informationen zum Unternehmen

Name des Unternehmens _____

Größe des Unternehmens
1 - 19 Mitarbeiter
20 - 99 Mitarbeiter
100 - 499 Mitarbeiter
500 - 999 Mitarbeiter
1.000 - 4.999 Mitarbeiter
mehr als 5.000 Mitarbeiter

Webseite des Unternehmens ___

Land _____ Stadt _____

Namen des Befragten ___ E-Mail-Adresse ___

Industriebereich

- ☐ Luft- und Raumfahrt, Verteidigung und innere Sicherheit
- ☐ Automobilindustrie
- ☐ Finanz- und Kapitalmärkte
- ☐ Chemieindustrie
- ☐ Kommunikationsindustrie
- ☐ Energie, Versorgung und Bergbau
- ☐ Ingenieurs- und Bauwesen
- ☐ Unterhaltung und Medien
- ☐ Finanzdienstleister
- ☐ Forstwirtschaft, Papier- und Verpackungsindustrie
- ☐ öffentlicher Dienst
- ☐ Gesundheitswesen
- ☐ Gastronomie und Hotelwesen
- ☐ industrielle Fertigung
- ☐ Versicherungswesen
- ☐ Pharmaindustrie und Biowissenschaften
- ☐ Einzelhandel
- ☐ Technologie und Entwicklung
- ☐ Transport und Logistik
- ☐ Sonstiges (bitte angeben): ___

Wie kann ein Experte der digitalen Linguistik dazu beitragen, Ihr Unternehmen auszubauen?

Muss Ihr Unternehmen die **Inhalte** folgender Textsorten ordnen, analysieren oder kategorisieren?

- **E-Mails** und andere Arten schriftlicher Kommunikation
  ☐ Wird bereits umgesetzt ☐ In Planung ☐ Nicht sicher ☐ Nein

- Support- und/oder Kunden**anfragen**
  ☐ Wird bereits umgesetzt ☐ In Planung ☐ Nicht sicher ☐ Nein

- **Dokumente** (Wissensdatenbanken, interne Dokumente, Verträge, rechtswirksame Dokumente)
  ☐ Wird bereits umgesetzt ☐ In Planung ☐ Nicht sicher ☐ Nein

- Ihre eigenen **Webseiten**, **sozialen Medien**, Ihre **Werbe**-E-Mails oder die Ihrer Konkurrenz
  ☐ Wird bereits umgesetzt ☐ In Planung ☐ Nicht sicher ☐ Nein

Besteht Bedarf, komplexe Textsuchen durchzuführen, d.h. große Textmengen, bei denen ähnliche oder verwandte Wörter und ihre verschiedenen Wortformen miteingeschlossen werden, schnell und intelligent zu durchsuchen?

Bedarf	☐ Jetzt ☐ In Zukunft ☐ Nicht sicher ☐ Nein
Outsourcen *)	☐ Ja ☐ Vielleicht ☐ Nein, wird inhouse erledigt
Mehrsprachig **)	☐ Ja ☐ Vielleicht ☐ Nein

Was machen Sie mit diesen Daten? Wer führt diese Aufgaben aus?

1. **Kategorisieren** nach Themen, **Markieren** oder **Kennzeichnen**

Wann?	☐ Wird bereits umgesetzt ☐ In Zukunft ☐ Nicht sicher ☐ Nein
Outsourcen *)	☐ Ja ☐ Vielleicht oder teilweise ☐ Nein, wird in-House erledigt
Mehrsprachig **)	☐ Ja ☐ Vielleicht ☐ Nein

2. Analysieren der **Grundstimmung** in Blog-Posts, Kommentaren, Beiträgen in sozialen Medien und Medien im Allgemeinen

Wann?	☐ Wird bereits umgesetzt ☐ in Zukunft ☐ nicht sicher ☐ Nein
Outsourcen *)	☐ Ja ☐ Vielleicht oder teilweise ☐ Nein, wird in-House erledigt
Mehrsprachig **)	☐ Ja ☐ Vielleicht ☐ Nein

3. Überprüfen der Konsistenz von **Terminologie** oder von **kontrollierter Sprache** in Texten oder Datenbanken

Wann?	☐ Wird bereits umgesetzt ☐ In Zukunft ☐ Nicht sicher ☐ Nein
Outsourcen *)	☐ Ja ☐ Vielleicht oder teilweise ☐ Nein, wird in-House erledigt
Mehrsprachig **)	☐ Ja ☐ Vielleicht ☐ Nein

Auflistung der Unterpunkte in der Reihenfolge ihrer Priorität (Zahlen durch Komma trennen) ___

Haben Sie Kommentare zu den Fragen auf dieser Seite?

_____

*) Würden Sie Ihr eigenes System entwickeln und/oder Ihr eigenes Personal einsetzen, um das Ziel zu erreichen, oder würden Sie eine fertige Lösung bevorzugen?
**) Sollte die Technologie in mehr als einer Sprache vorliegen?

Brauchen Sie Personen in Ihrem Team mit Know-how in **automatischen Kommunikationstechnologien** wie z.B.:

1. automatisches, thematisch-passendes **Antworten** auf E-Mails oder andere Mitteilungen (möglicherweise vor der Übergabe an das Call Center)
   Bedarf ☐ Jetzt ☐ In Zukunft ☐ Nicht sicher / Nein
   Outsourcen *) ☐ Ja ☐ Vielleicht ☐ Nein, wird in-House erledigt
   Mehrsprachig **) ☐ Ja ☐ Vielleicht ☐ Nein

2. **Chatbots** für automatisierte Chat-Kommunikation auf Webseiten (möglicherweise nur, um einige grundlegende Informationen vor der Übergabe an das Call Center zu erhalten)
   Bedarf ☐ Jetzt ☐ In Zukunft ☐ Nicht sicher / Nein
   Outsourcen *) ☐ Ja ☐ Vielleicht oder teilweise ☐ Nein, wird in-House erledigt
   Mehrsprachig **) ☐ Ja ☐ Vielleicht ☐ Nein

3. automatisierte **Sprache/Dialoge** für die Telefon-Kommunikation mit Kunden
   Bedarf ☐ Jetzt ☐ In Zukunft ☐ Nicht sicher / Nein
   Outsourcen *) ☐ Ja ☐ Vielleicht oder teilweise ☐ Nein, wird in-House erledigt
   Mehrsprachig **) ☐ Ja ☐ Vielleicht ☐ Nein

4. animierte digitale Figuren (Avatare) auf Webseiten
   Bedarf ☐ Jetzt ☐ In Zukunft ☐ Nicht sicher / Nein
   Outsourcen *) ☐ Ja ☐ Vielleicht oder teilweise ☐ Nein, wird in-House erledigt
   Mehrsprachig **) ☐ Ja ☐ Vielleicht ☐ Nein

Auflistung der Unterpunkte in der Reihenfolge ihrer Priorität (Zahlen durch Komma trennen) ___

Wie nutzen Ihre seh- oder hörbeeinträchtigten **Mitarbeiter** Ihre Systeme und wie interagieren sie mit Ihrer Web-Präsenz?

_____

Wie interagieren Ihre seh- oder hörbeeinträchtigten **Kunden** mit Ihrer Web-Präsenz?

_____

Haben Sie Kommentare zu den Fragen auf dieser Seite?

_____

*) Würden Sie Ihr eigenes System entwickeln und/oder Ihr eigenes Personal einsetzen, um das Ziel zu erreichen, oder würden Sie eine fertige Lösung bevorzugen?
**) Sollte die Technologie in mehr als einer Sprache vorliegen?

Werden in Ihrem Unternehmen regelmäßig Personen oder Tools eingesetzt, um ...

1. Übersetzungen anzufertigen?
   Bedarf ☐Jetzt ☐In Zukunft ☐Nicht sicher, vielleicht nicht
   digitale Werkzeuge ☐Werden bereits genutzt ☐Werden noch nicht verwendet
   Outsourcen *) ☐Ja ☐Nein
   [wenn Bedarf: jetzt]
   In wie viele Sprachen wird bei Ihnen regelmäßig übersetzt?
   Wie viele Seiten werden pro Woche geschätzt übersetzt?
   [wenn digitale Werkzeuge: werden bereits genutzt]
   Welche Tools verwenden Sie:
   Übersetzungs-Tools (CAT-Tools)
   Maschinelle Übersetzungssysteme
   Sonstiges (bitte angeben): ___
   [Wenn outsourcen: nein]
   Wie viel Personal beschäftigt Ihr Unternehmen für Übersetzungsaufgaben?

2. Webinhalte zu erstellen?
   Bedarf ☐Jetzt ☐In Zukunft ☐Nicht sicher, vielleicht nicht
   Outsourcen *) ☐Ja ☐Vielleicht oder teilweise ☐Nein, wird in-House erledigt
   Mehrsprachig **) ☐Ja ☐Vielleicht ☐Nein

3. Werbetexte zu erstellen?
   Bedarf ☐Jetzt ☐in Zukunft ☐Nicht sicher, vielleicht nicht
   Outsourcen *) ☐Ja ☐Vielleicht oder teilweise ☐Nein, wird in-House erledigt
   Mehrsprachig **) ☐Ja ☐Vielleicht ☐Nein

Auflistung der Unterpunkte in der Reihenfolge ihrer Priorität ___

Nutzen Sie manchmal **automatische Schreibhilfen**? (Tools, die Feedback geben, während der Benutzer tippt, und die z. B. Formulierungsvorschläge anbieten oder auf die korrekte Verwendung von Terminologie auf Basis der Unternehmensrichtlinien hinweisen.)
Bedarf ☐Jetzt ☐In Zukunft ☐Nicht sicher, vielleicht auch nicht
Outsourcen *) ☐Ja ☐Vielleicht oder teilweise ☐Nein, wird in-House erledigt
Mehrsprachig **) ☐Ja ☐Vielleicht ☐Nein

Haben Sie Kommentare zu den Fragen auf dieser Seite?

_____

*) Würden Sie Ihr eigenes System entwickeln und/oder Ihr eigenes Personal einsetzen, um das Ziel zu erreichen, oder würden Sie eine fertige Lösung bevorzugen?
**) Sollte die Technologie in mehr als einer Sprache vorliegen?

Pomagajte nam pri izobraževanju digitalnih strokovnjakov

O projektu DigiLing

www.digiling.eu

DigiLing je projekt, financiran iz programa Erasmus+, katerega cilj je zadovoljiti rastoče potrebe po strokovnjakih s področij obdelave, analize, iskanja in uporabe velikih količin besedilnih podatkov, kot so e-pošta, uporabniška podpora, interni dokumenti ter besedila s spletnih strani, družabnih omrežij in drugih medijev. V okviru projekta bodo izdelana orodja za e-učenje temeljnih veščin in znanj digitalnega jezikoslovja, namenjena izobraževanju **digitalnih jezikovnih strokovnjakov** za podjetja, kot je vaše.

S pomočjo **digitalnega jezikovnega strokovnjaka** bo vaše podjetje bolj prepoznavno, učinkoviteje bo nagovarjalo vaše ciljne skupine ter bolj poglobljeno razumelo lokalne in mednarodne trende ter strategije konkurenčnih podjetij. Z drugimi besedami, digitalni jezikovni strokovnjak vam bo omogočil izluščiti kar največ iz vaših obstoječih besedilnih podatkov.

S pričujočo raziskavo želimo ugotoviti, v kakšnem obsegu naštete aktivnosti že izvajate oziroma kako koristen bi za vas bil digitalni jezikovni strokovnjak. Tako bomo lahko natančneje izdelali učne vsebine za digitalne jezikovne strokovnjake, ki bodo posredovale znanja, po katerih je na trgu največ povpraševanja.

Podatki o podjetju

Ime podjetja _____

Število zaposlenih:
1–19
20–99
100–499
500–999
1000–4999
več kot 5000

Spletna stran podjetja _____

Država _____ Mesto _____

Ime anketiranca (neobvezno) _____ Elektronski naslov_____

Gospodarska panoga

☐ Letalska, obrambna in varnostna industrija
☐ Avtomobilska industrija
☐ Bančništvo in kapitalski trgi
☐ Kemična industrija
☐ Komunikacije
☐ Energetska industrija, rudarstvo, komunala
☐ Inženiring in gradbeništvo
☐ Razvedrilo in mediji
☐ Finančne storitve
☐ Lesna, papirna in embalažna industrija
☐ Država in javne storitve
☐ Zdravstvo
☐ Turizem in prosti čas
☐ Industrijska proizvodnja
☐ Zavarovalništvo
☐ Farmacija & biološke znanosti
☐ Maloprodaja in trgovina
☐ Tehnologija
☐ Transport in logistika
☐ Drugo (prosimo, navedite)_____

Kako lahko digitalni jezikovni strokovnjak pomaga vašemu podjetju?

Ali vaše podjetje potrebuje razvrščanje in analizo dokumentov **glede na vsebino** v katerem od naslednjih primerov?

- **e-pošta** in druga **pisna komunikacija**
  ☐ že izvajamo ☐ načrtujemo ☐ nismo prepričani ☐ ne

- **uporabniška podpora** (support tickets), **odgovarjanje na vprašanja** strank
  ☐ že izvajamo ☐ načrtujemo ☐ nismo prepričani ☐ ne

- **dokumenti** (podatkovne baze, interni dokumenti, pogodbe, pravni dokumenti)
  ☐ že izvajamo ☐ načrtujemo ☐ nismo prepričani ☐ ne

- vaše ali konkurenčne **spletne strani**, **družabna omrežja** ali **promocijska** e-pošta
  ☐ že izvajamo ☐ načrtujemo ☐ nismo prepričani ☐ ne

Ali potrebujete napredno besedilno iskanje, tj. hitro in pametno **iskanje** po velikih količinah besedil, ki zna najti tudi podobne ali sorodne besede ter njihove pregibne oblike?

potrebujemo	☐ zdaj	☐ v prihodnosti	☐ nismo prepričani	☐ ne
zunanji izvajalci*	☐ vsekakor	☐ mogoče	☐ ne, interno	
v več jezikih**	☐ da	☐ mogoče	☐ ne	

Kaj počnete s podatki? Kdo izvaja ta opravila?

1. **razvrščanje** po temah, **označevanje** ali **opremljanje s ključnimi besedami**

kdaj?	☐ že izvajamo	☐ v prihodnosti	☐ nismo prepričani	☐ ne
zunanji izvajalci*	☐ vsekakor	☐ mogoče ali delno	☐ ne, interno	
v več jezikih**	☐ da	☐ mogoče	☐ ne	

2. analiza **stališč** v medijih, spletnih dnevnikih, komentarjih in objavah na družabnih omrežjih

kdaj?	☐ že izvajamo	☐ v prihodnosti	☐ nismo prepričani	☐ ne
zunanji izvajalci*	☐ vsekakor	☐ mogoče ali delno	☐ ne, interno	
v več jezikih**	☐ da	☐ mogoče	☐ ne	

3. preverjanje **terminološke** doslednosti ali uporaba **nadzorovanega jezika** v besedilih in podatkovnih bazah

kdaj?	☐ že izvajamo	☐ v prihodnosti	☐ nismo prepričani	☐ ne
zunanji izvajalci*	☐ vsekakor	☐ mogoče ali delno	☐ ne, interno	
v več jezikih**	☐ da	☐ mogoče	☐ ne	

Prosimo, razvrstite številke po pomembnosti (ločite z vejicami): _____

Imate kakšne komentarje na zgornja vprašanja? _____

* Bi raje izdelali lasten sistem ter zaposlili lastnega strokovnjaka ali najeli zunanjega izvajalca?

** Ali bi ta tehnologija morala delovati z več kot enim jezikom?

Ali potrebujete strokovnjaka, ki obvlada tehnologije za samodejno komunikacijo, kot npr.:

1. samodejno pametno **odgovarjanje** na e-pošto in druga komunikacija (denimo preden stranko prevzame svetovalec)

potrebujemo	☐ zdaj	☐ v prihodnosti	☐ nismo prepričani, ne
zunanji izvajalci*	☐ vsekakor	☐ mogoče	☐ ne, interno
v več jezikih**	☐ da	☐ mogoče	☐ ne

2. **robotski sogovornik** za samodejno komunikacijo z obiskovalci spletne strani (denimo za podajanje osnovnih informacij, preden pogovor prevzame svetovalec)

potrebujemo	☐ zdaj	☐ v prihodnosti	☐ nismo prepričani, ne
zunanji izvajalci*	☐ vsekakor	☐ mogoče	☐ ne, interno
v več jezikih**	☐ da	☐ mogoče	☐ ne

3. sintetiziran **govor / pogovor** za telefonsko komunikacijo s strankami

potrebujemo	☐ zdaj	☐ v prihodnosti	☐ nismo prepričani, ne
zunanji izvajalci*	☐ vsekakor	☐ mogoče	☐ ne, interno
v več jezikih**	☐ da	☐ mogoče	☐ ne

4. animirani agenti (avatarji) na spletnih straneh

potrebujemo	☐ zdaj	☐ v prihodnosti	☐ nismo prepričani, ne
zunanji izvajalci*	☐ vsekakor	☐ mogoče	☐ ne, interno
v več jezikih**	☐ da	☐ mogoče	☐ ne

Prosimo, razvrstite številke po pomembnosti (ločite z vejicami): _____

Na kakšen način lahko **zaposleni** z okvarami vida ali sluha uporabljajo vaše sisteme in spletne strani?

_____

Na kakšen način lahko **stranke** z okvarami vida ali sluha uporabljajo vaše spletne strani?

_____

Imate kakšne komentarje na zgornja vprašanja? _____

* Bi raje izdelali lasten sistem ter zaposlili lastnega strokovnjaka ali najeli zunanjega izvajalca?

** Ali bi ta tehnologija morala delovati z več kot enim jezikom?

Ali vaše podjetje redno potrebuje osebje in/ali orodja za naslednja opravila?

1. prevajanje

potrebujemo	☐ zdaj	☐ v prihodnosti	☐ nismo prepričani, morda ne
računalniška orodja	☐ uporabljamo	☐ ne uporabljamo	
zunanji izvajalci*	☐ da	☐ ne	

[če ste odgovorili s "potrebujemo zdaj"]
V koliko jezikov prevajate redno?
Približno koliko strani na teden?
[če uporabljate računalniška orodja]
Katera orodja so to?
☐ računalniško podprto prevajanje (orodja CAT)
☐ strojno prevajanje
☐ drugo (prosimo, navedite):_____
*[če prevajanja **ne** oddajate zunanjim izvajalcem]*
Koliko oseb, ki so zadolžene za prevajanje, zaposlujete?

2. ustvarjanje spletnih vsebin

potrebujemo	☐ zdaj	☐ v prihodnosti	☐ nismo prepričani, morda ne
zunanji izvajalci*	☐ vsekakor	☐ mogoče ali delno	☐ ne, interno
v več jezikih**	☐ da	☐ mogoče	☐ ne

3. ustvarjanje oglasnih in promocijskih vsebin

potrebujemo	☐ zdaj	☐ v prihodnosti	☐ nismo prepričani, morda ne
zunanji izvajalci*	☐ vsekakor	☐ mogoče ali delno	☐ ne, interno
v več jezikih**	☐ da	☐ mogoče	☐ ne

Prosimo, razvrstite številke po pomembnosti (ločite z vejicami): _____

Ali potrebujete **digitalnega pomočnika pri pisanju (writing assistant)** – orodje, ki med pisanjem sproti ponuja predloge za lepše izražanje ter bdi nad pravilno uporabo terminologije v skladu s standardi vašega podjetja?

potrebujemo	☐ zdaj	☐ v prihodnosti	☐ nismo prepričani, morda ne
zunanji izvajalci*	☐ vsekakor	☐ mogoče ali delno	☐ ne, interno
v več jezikih**	☐ da	☐ morda	☐ ne

Imate kakšne komentarje na zgornja vprašanja? _____

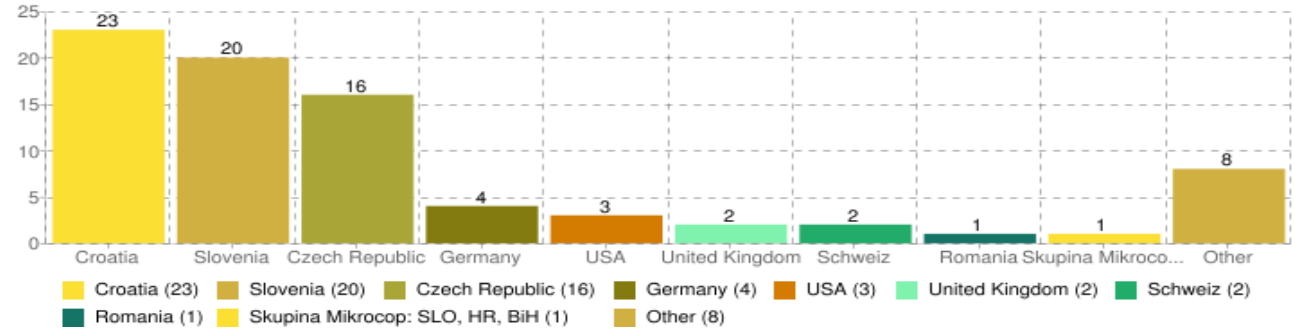
* Bi raje izdelali lasten sistem ter zaposlili lastnega strokovnjaka ali najeli zunanjega izvajalca?
** Ali bi ta tehnologija morala delovati z več kot enim jezikom?

Appendix 4: Visual report

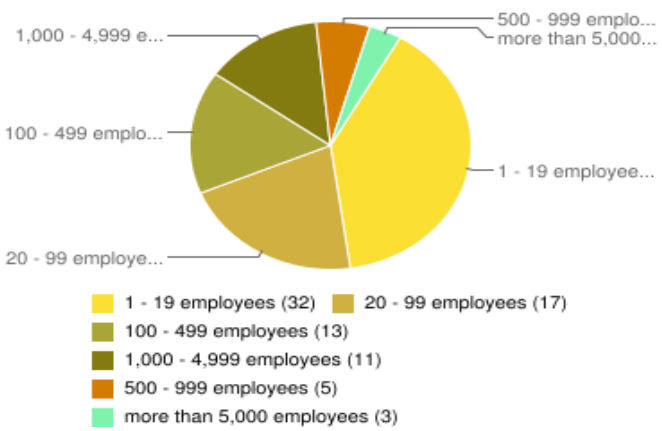
pages 1 – 5

of the online form (as paginated when viewed on the screen)

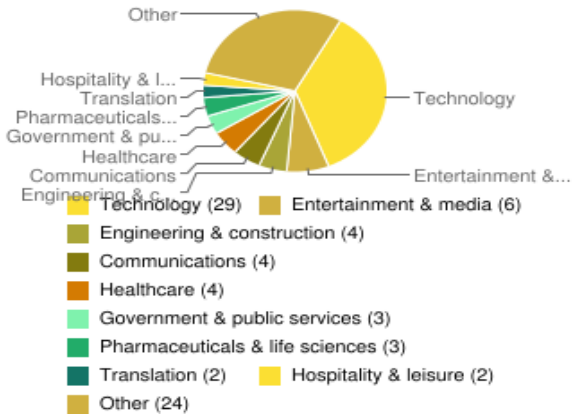
country



company size



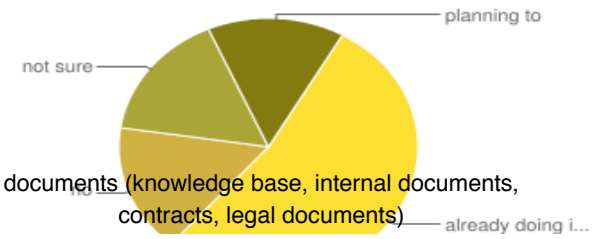
Industry



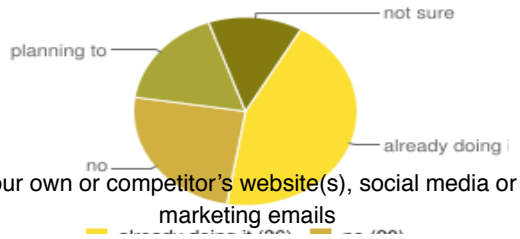
Does your company need to sort, analyse or categorize any of these **by the content**?

emails and other written communication

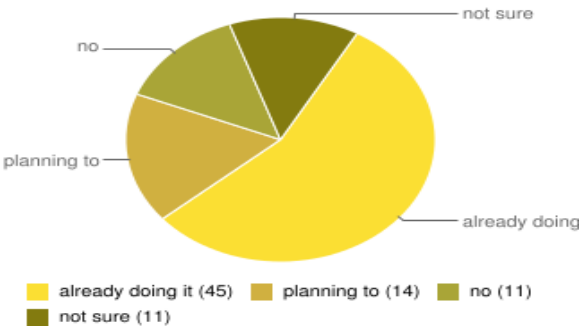
support tickets, customer queries

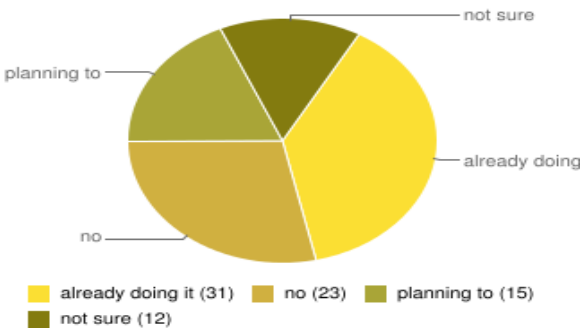


documents (knowledge base, internal documents, contracts, legal documents)



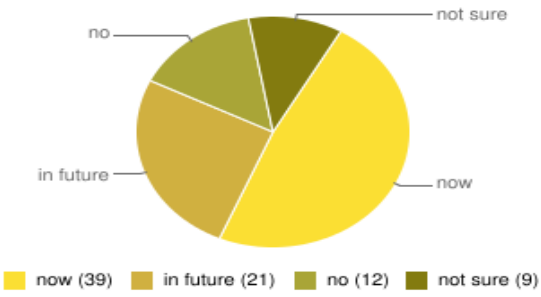
your own or competitor's website(s), social media or marketing emails



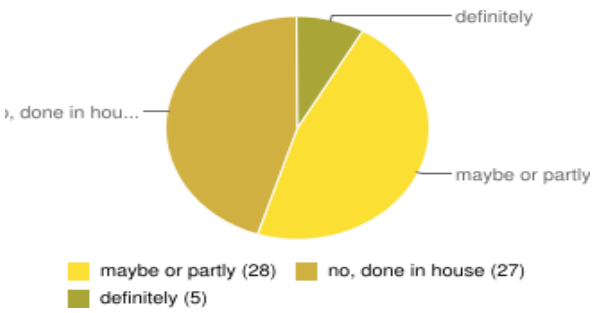


Is there a need for sophisticated **text searching**, i.e. fast intelligent searching of large amounts of text able to include similar or related words and their different word forms.

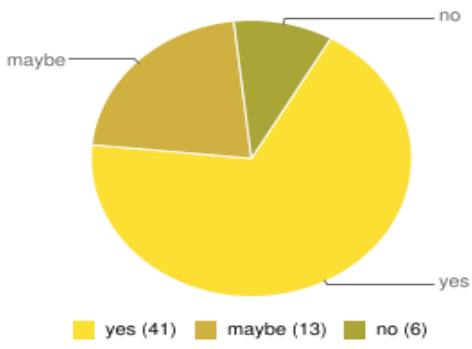
need



outsource



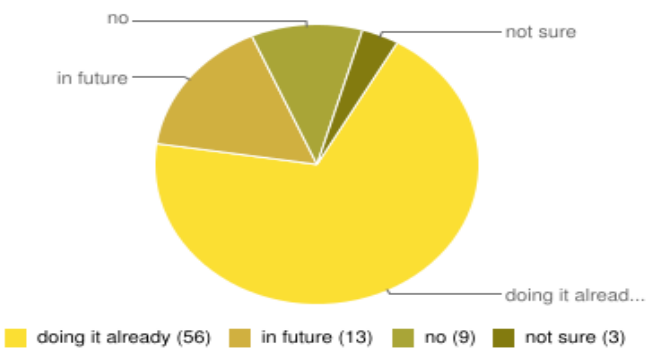
multi-lingual



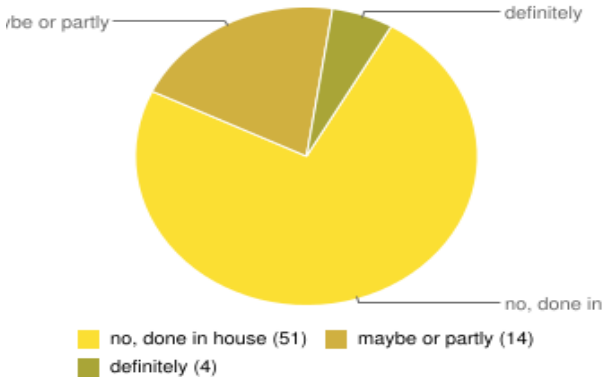
What do you do with the data? Who performs these tasks?

1 **categorization** by topics, **labelling** or **tagging**

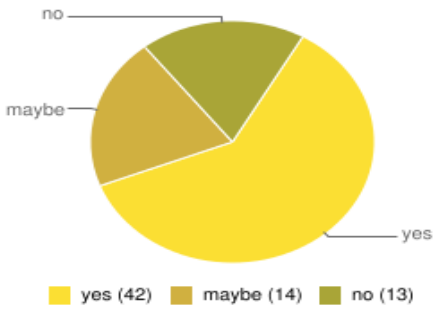
when?



outsource

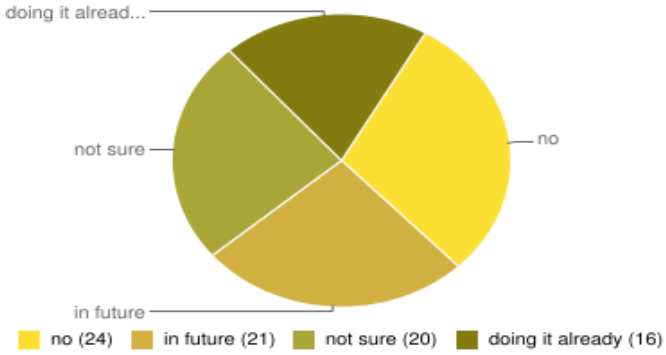


multi-lingual

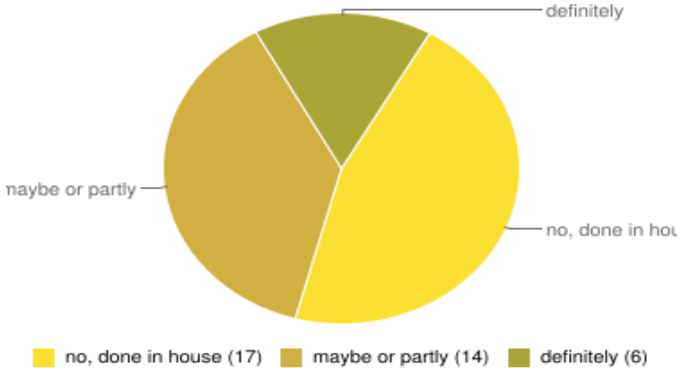


2 analysis of **attitudes** in blog posts, comments and social media posts and media in general

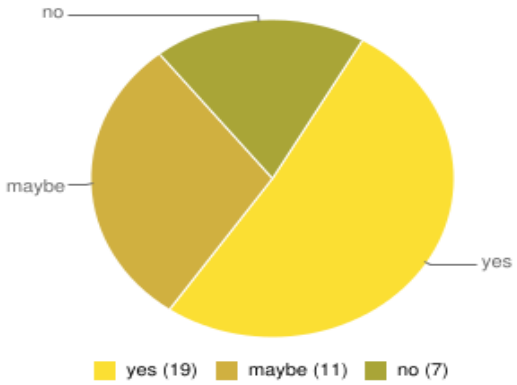
when?



outsource

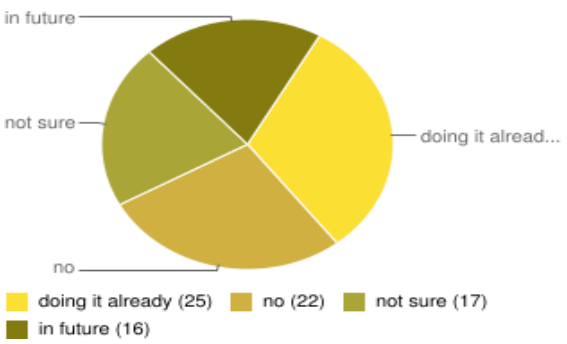


multi-lingual

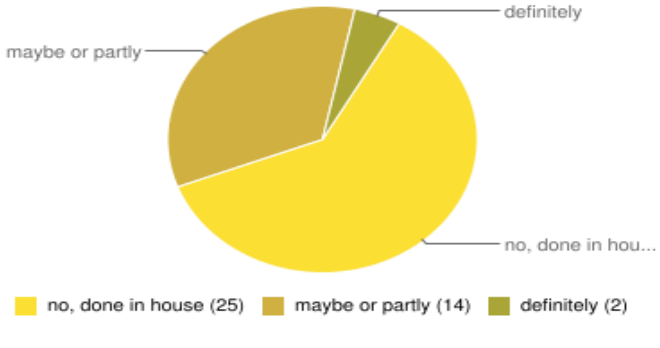


3 consistency check of **terminology** or controlled **language** used in texts or databases

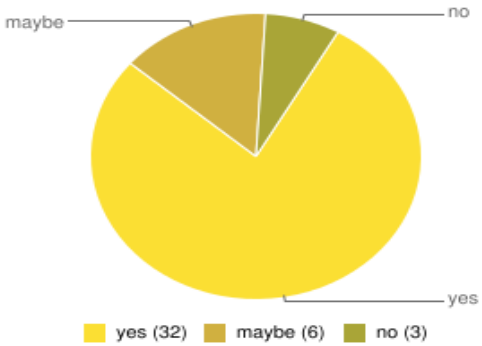
when?



outsource



multi-lingual



company name	order of	company name	Any comments on the three...
Jumpshot	321	Termitnjak d.o.o	
Termitnjak d.o.o	312	Sandberg Translation Partners Ltd	
Leksikografski zavod Miroslav Krleža	312		

company name	order of	company name	Any comments on the three...
Xellia	312	Modra zavarovalnica, d.d.	medijsko spremljamo le besedila v slovenskem jeziku, v tujih jezikih pride v poštev evropska zakonodaja (slabi ali dvomljivi prevodi v slovenščino), analiza primerjalnih podjetij v EU, dokumentacija o naložbah in investicijah v angleščini, ki jih dobimo od tujih finančnih posrednikov...
Marko G.	312		
ZOP-CR d.o.o.	312		
		Mikrocop d.o.o.	

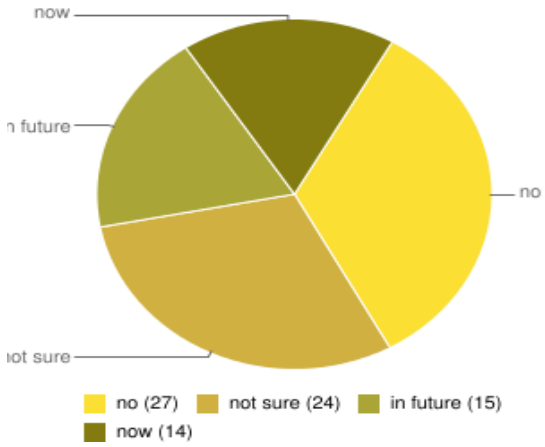
pages 6 - 11

of the online form (as paginated when viewed on the screen)

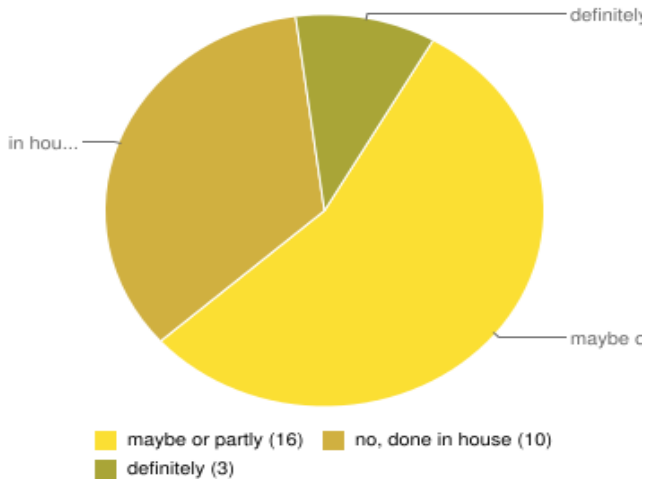
Do you need people with expertise in **automated communication** technologies such as:

1 automatic relevant **responses** to emails or other communication (possibly before handing over to a live agent)

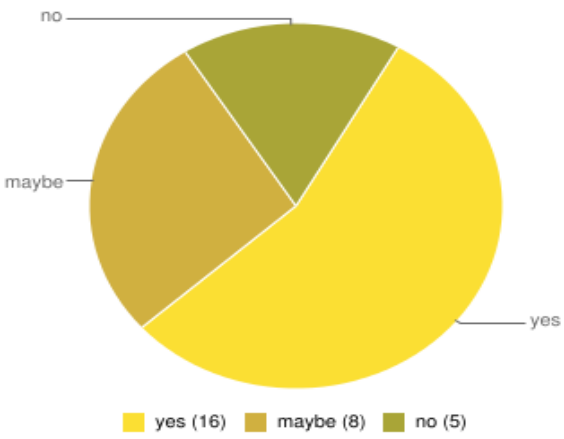
need



outsource



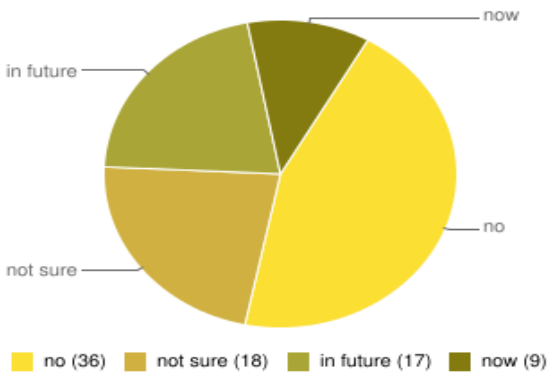
multi-lingual

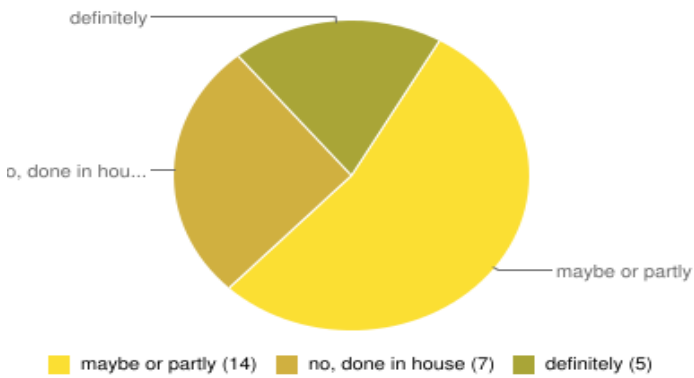


2 **chat bots** for automated website chat communication (possibly to only get some basic information before handing over to a live agent)

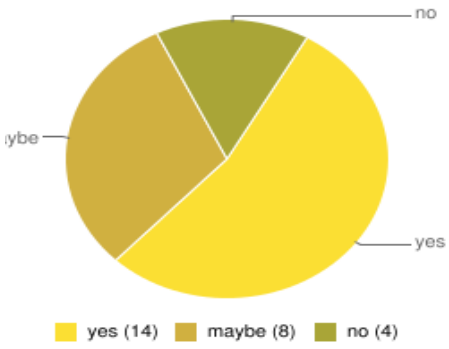
need

outsource



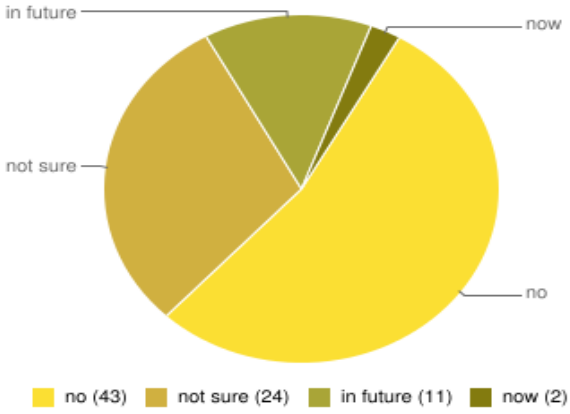


multi-lingual

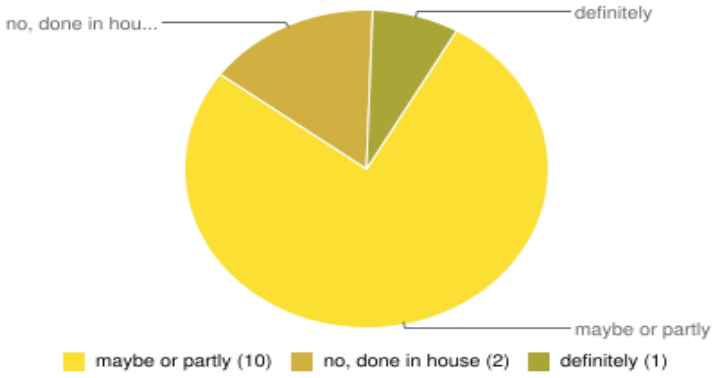


3 automated **speech / dialogue** for phone communications with clients

need

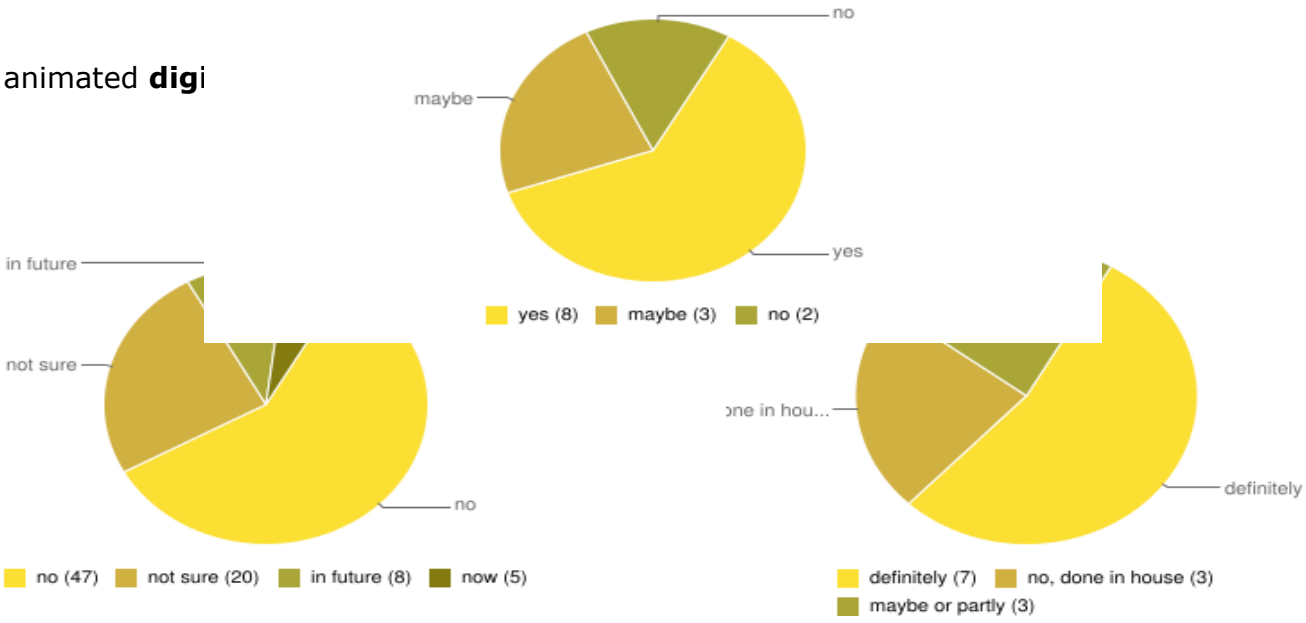


outsource

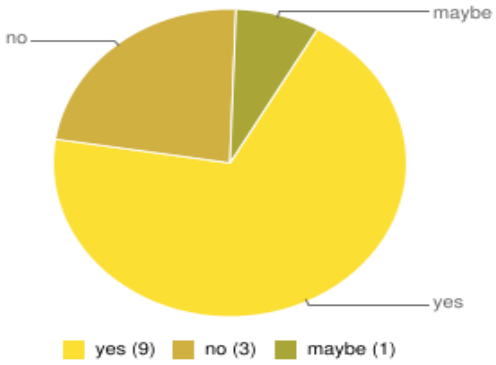


multi-lingual

4 animated **digi**



multi-lingual



company name	order of ir
Jumpshot	1
Termitnjak d.o.o	2314
Leksikografski zavod Miroslav Krleža	1243
Xellia	9
Marko G.	1324
ZOP-CR d.o.o.	1234
ACREA CR	1324

company name	Any comments on the three...
Termitnjak d.o.o	
Sandberg Translation Partners Ltd	
Modra zavarovalnica, d.d.	
Mikrocop d.o.o.	
AXIS, Projektne IT rešitve in storitve, d.o.o.	
Dnevnik, d.d.	Nisem prepričan, da razumem, kaj mislite z animiranimi agenti.
Delo d.o.o.	

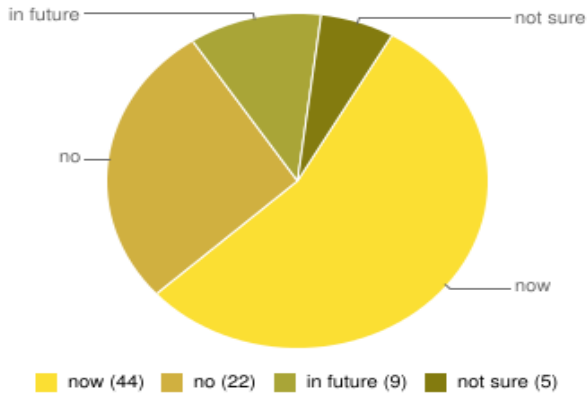
How do your visually/hear...	How do your visually/hea...
-	-
Virtually all of our internal communication is written. We do not have any visually impaired employees for whom this would be a problem. We have a hearing impaired employee, and this text-focused method suits him well.	Only a handful of our employees needs to interact without web pre None of them are visually/hearing impaired.
nimamo takih zaposlenih	uporaba lupe za slabovidne, za slepe in gluhe nimamo posebne p

How do your visually/hear...	How do your visually/hea...
/	/
Okvara vida - ni omogočeno. Okvara sluha - ni ovir.	Okvara vida - ni omogočeno. Okvara sluha - ni ovir.
	Ker smo v osnovi tekstovno naravnani, za stranke z okvarami sluh posebnih težav, za tiste z (nepopolnimi) okvarami vida pa pride v

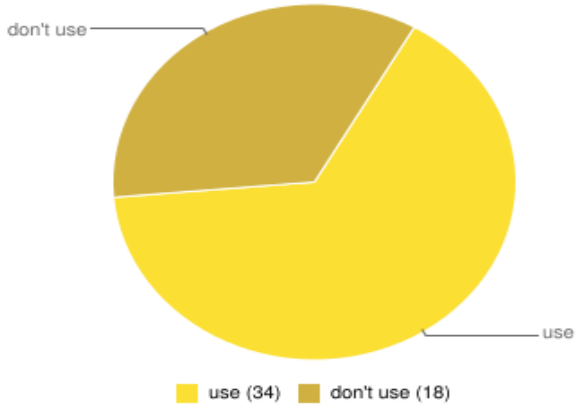
Does your company regularly need people and/or tools for…?

1 translation

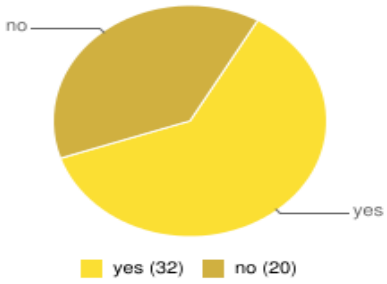
need



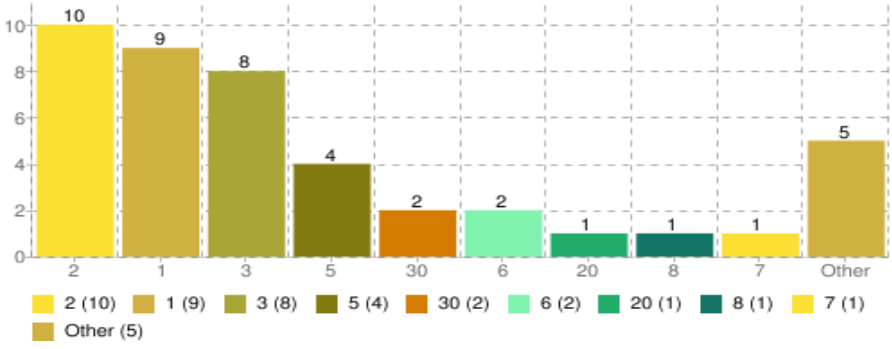
digital tools



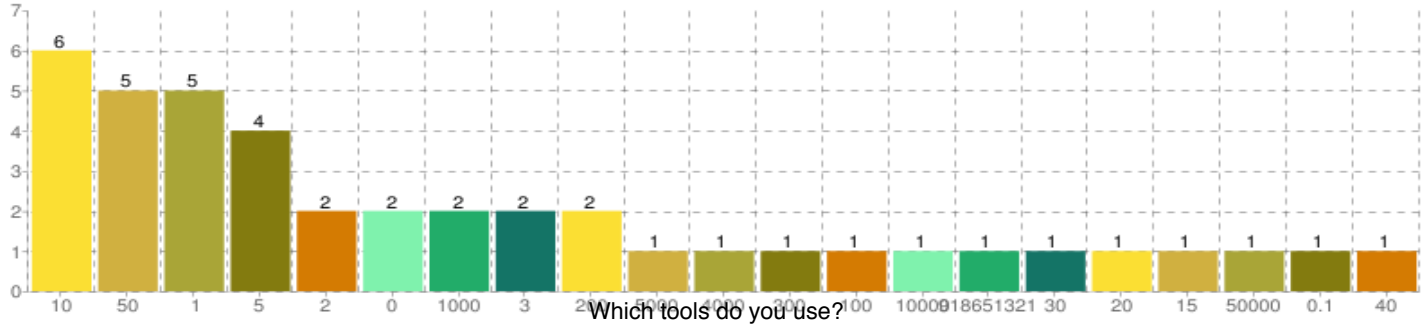
outsource



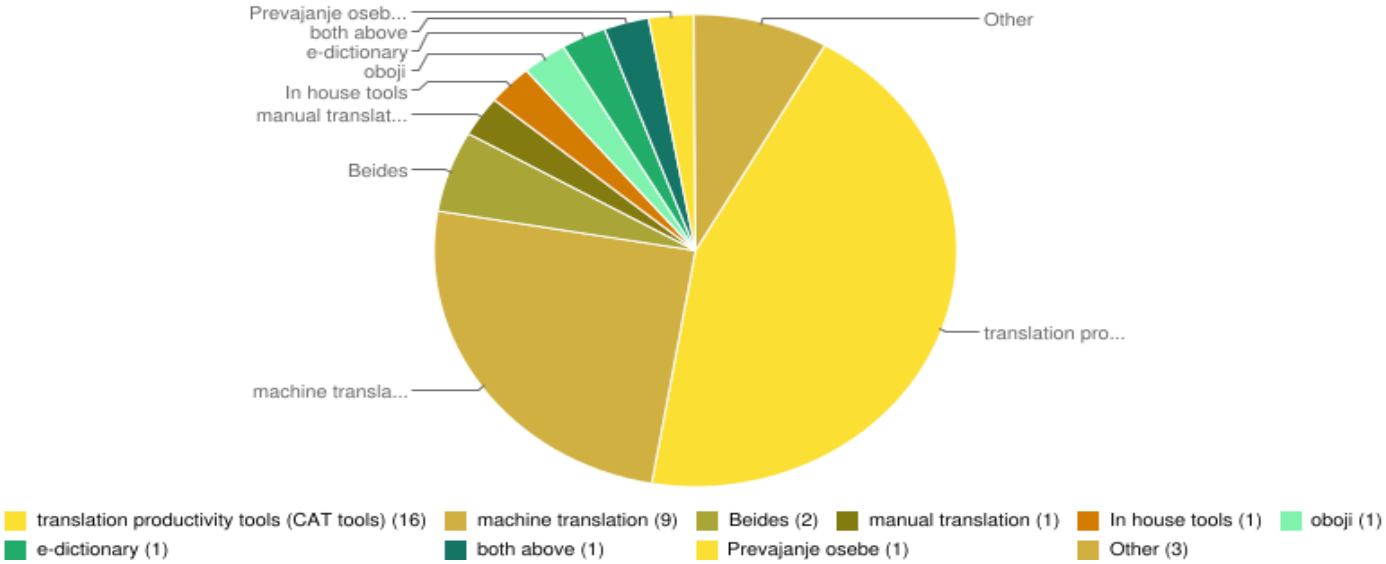
How many languages do you regularly translate into?



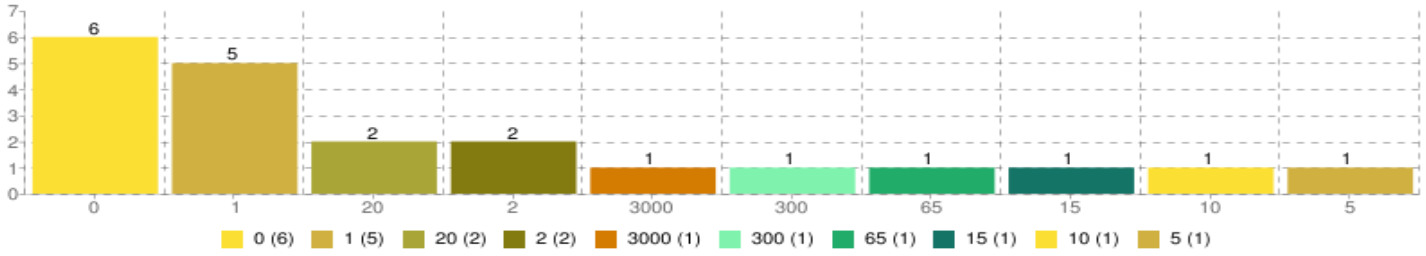
Estimated volume of pages per week?



Which tools do you use?



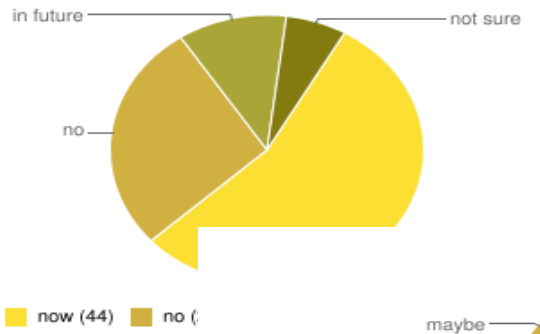
How many persons does your company employ for translation tasks?

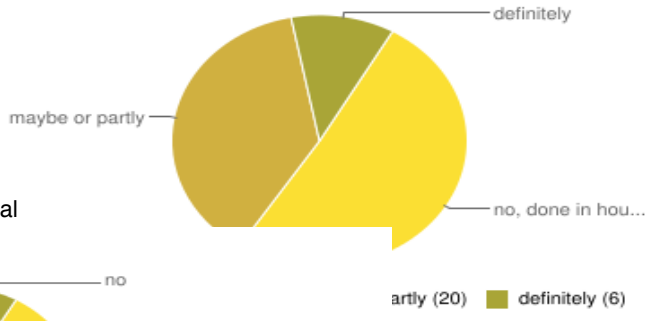


2 web content authoring

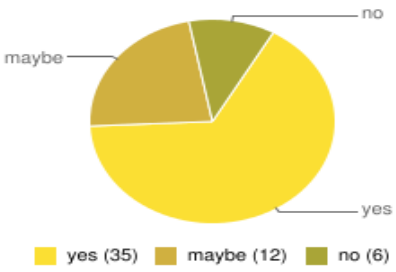
need

outsource



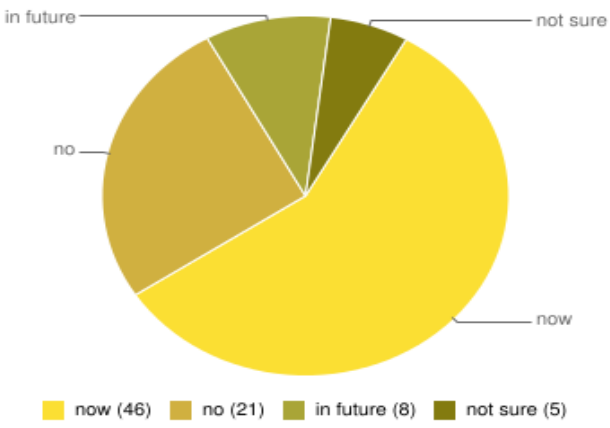


multi-lingual

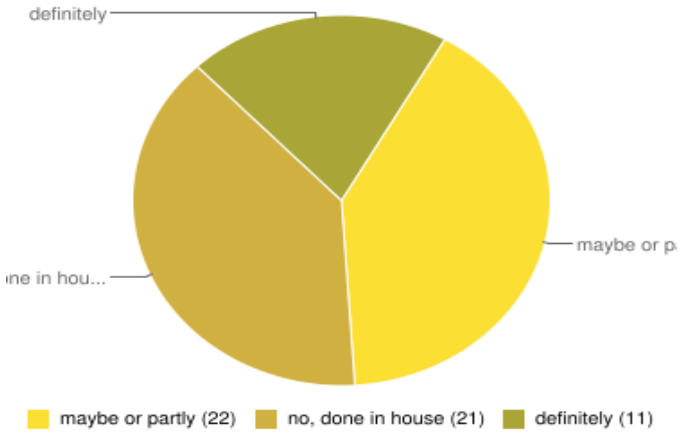


3 copywriting

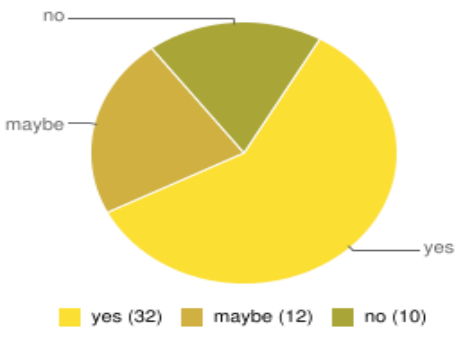
need



outsource



multi-lingual

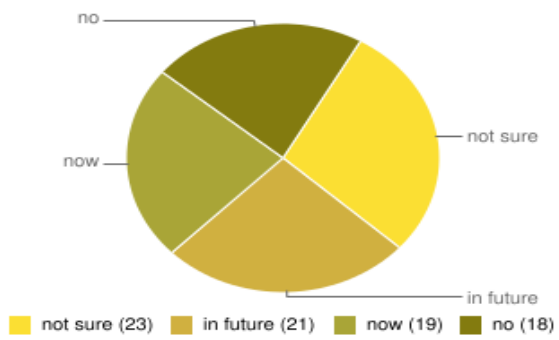


company name	order of in	company name	Any comments on the three...
Termitnjak d.o.o	213	Termitnjak d.o.o	

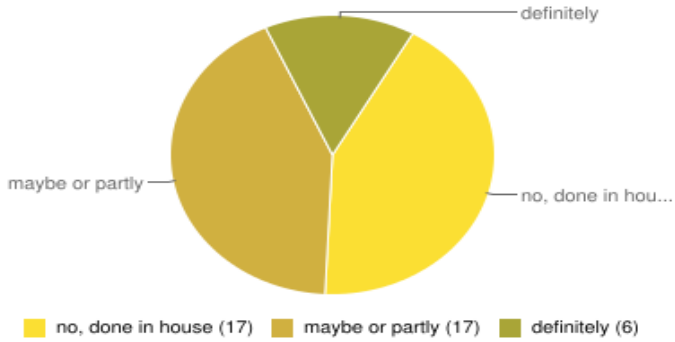
company name	order of in	company name	Any comments on the three...
Sandberg Translation Partners Ltd	123	Sandberg Translation Partners Ltd	In your question 1 about Translation, I was not able to select both CAT tools and MT. So I left MT empty even though we use it as well.
Modra zavarovalnica, d.d.	123	Modra zavarovalnica, d.d.	naše stranke so zgolj Slovenci.
Mikrocop d.o.o.	321	Mikrocop d.o.o.	
AXIS, Projektne IT rešitve in storitve, d.o.o.	123	AXIS, Projektne IT rešitve in storitve, d.o.o.	

Is there a use case for **writing assistants** - a tool providing feedback as the user types and offering suggestions for better wording or correct use of terminology based on the company standards?

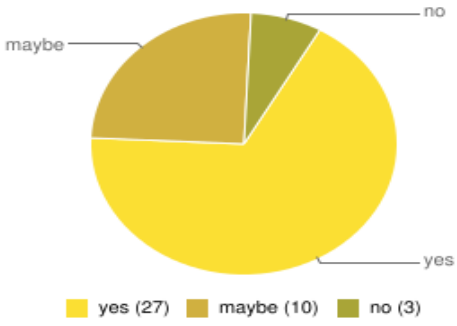
need



outsource



multi-lingual



Any comments on the topic above?

Showing 7 responses

Relatively distant future
26/01/2017 4:00 PM

ne
8/02/2017 2:40 PM

Pametno.
8/02/2017 9:53 PM

This is a great idea, but I'd like to see a demo first.
13/02/2017 4:07 PM

Uporabljam Grammarly: https://app.grammarly.com
14/02/2017 1:11 PM

ne razumem dobro vprašanj, vse od tega že obstaja in uporabljamo.
14/02/2017 4:34 PM

Not exactly, but keen on writing assistant to users to properly write answers to clinical questions!
8/03/2017 12:22 PM