Language data and project specialist
A new modular profile for graduates in language-related disciplines

UPSKILLS Intellectual output 1.6

Compiled by:

Maja Miličević Petrović[*], Silvia Bernardini[*], Adriano Ferraresi[*],
Gaia Aragrande[*], Alberto Barrón-Cedeño[*]

[*] University of Bologna

UPSKILLS Consortium:

Table of contents

Executive Summary

The UPSKILLS needs analysis explored the current academic offer in language- and linguistics-related fields (modern languages and cultures, translation, general linguistics, etc.) and the requirements the job market has for graduates in these areas. The analysis highlighted the need for a new skill set and a new mind frame to meet the demands as well as the professional challenges of the industry. Taking into consideration the results of the individual components of the needs analysis, this final report outlines a new professional profile, that of the **language data and project specialist**, and includes a detailed description of the knowledge, skills and competences that present-day and future graduates in languages and linguistics should obtain to improve their employability in the digital business sector.

The report briefly summarises the outcomes of each component of the needs analysis. Starting from the survey of curricula, the need was showed to provide additional learning content in line with the requirements of the job market, and to enable students and educators to make the most out of the skills already present in their curricula. As concerns the survey of the literature, six skill clusters emerged as relevant, with a particular emphasis on research skills. The third component was based on a corpus of job advertisements that was perused for identifying typically required skills and competences, as well as typical job titles. In a cumulative process, the last two components built on the previous ones and led to the development of a questionnaire submitted to companies hiring linguists and language professionals. The skills and competences mentioned in the responses led in turn to the development of semi-structured interviews that were then carried out with industry representatives. All components generally confirmed that higher education curricula need to place more emphasis on the development of technical, managerial and transferable skills.

The profile of the language data and project specialist that emerges from this report is the result of a cumulative meta-analysis of the insights and outcomes of each component of the needs analysis. The profile is modular and adaptable to the needs of different study programmes and different positions in the job market. Thanks to this intrinsic flexibility in the concept of the profile itself, its educational outcomes revolve around two dimensions, a vertical one focused on seven main domains identified in the UPSKILLS needs analysis (disciplinary, (inter)cultural, technical, data-oriented, research-oriented, organisational and transversal), and a horizontal one based on the standard elements of learning outcomes (knowledge, skills and competences). Four sub-profiles are also defined based on whether the focus is more on research (language data analyst and language data scientist) or on management (language data manager and language project manager). In line with the objectives of UPSKILLS, the profile(s) will guide the creation of teaching and learning materials in three main domains (research skills, data acquisition skills and data handling skills), which will be the main focus of the other intellectual outputs of the project.

List of abbreviations

Abbreviation	Definition
AI	Artificial intelligence
BA	Bachelor of Arts
CAT	Computer-assisted translation
IT	Information technology
MA	Master of Arts
MSc	Master of Science
MT	Machine translation
NLP	Natural language processing
PhD	Doctor of Philosophy, Doctorate
STEM	Science, technology, engineering, and mathematics

List of Tables

1. Introduction

Throughout its history, the study of human language has taken many different forms, from the interpretation of literary texts to neuroimaging and computational linguistics. Despite the wide range of competences thus acquired, employment prospects for graduates in language-related disciplines (linguistics, modern languages, language pedagogy, translation and interpreting) tend to be rather limited and mostly focused on "traditional" occupations in teaching and translation, or neighbouring areas such as publishing. This stands in stark contrast with the potential employability of these graduates,[1] given the omnipresence of language and communication in society and the number of companies that make language one of their main businesses. Many companies across the world, from start-ups to technology giants, work with language data, and the demand for skills in language-related domains is constantly growing. A related demand is also surfacing in the public sector and in academic research, which growingly rely on empirical analyses of large-scale language data.

The UPSKILLS project team members share the view that the roots of the noted employment/employability paradox in the language professions largely lie in the mismatch between the type of knowledge and skills typically required by the contemporary job market and those traditionally included in early-to-mid stage university curricula (BAs and MAs). Language and linguistics sectors are being strongly affected by the changes brought about by the developments in the domain of technology and AI. These changes necessarily impact the job market, and even in the traditional language professions, there is an ever-growing need to embrace such developments as a new opportunity rather than with a fear of being replaced by automatic tools. The acceptance is already happening to some extent (many translation degrees, for example, include translation technologies in their curricula), but is not yet comprehensive. In addition, more recent AI-related changes that brought not just twists to the old occupations, but also entirely new options, seem not to be captured by higher education institutions, causing language and linguistics students to miss out on substantial opportunities. In particular, skills falling outside the disciplinary domain and the domain of technology appear to go under the radar and are consequently often neglected in degree programmes. Research skills can be taken as an illustrative example, still being rarely covered below the PhD level in language and linguistics degrees.

To address these issues, the UPSKILLS project aims to develop new learning and teaching materials that can be incorporated into existing curricula to enhance the skills identified as missing or requiring improvement. To detect such missing skills, a five-component needs analysis was conducted, consisting of: (1) a survey of language- and linguistics-related curricula at European universities, (2) a literature review, (3) an analysis of language industry and other related job advertisements, (4) a questionnaire-based survey of the industry needs, and (5) in-depth interviews with industry representatives. The five components

---

[1] Throughout the report, we use the term "employability" as also referring to freelancing opportunities.

aimed at providing a multi-faceted picture in which different points of view converge toward a comprehensive account of the current higher education offer in languages and linguistics, employers' needs, and institutional, academic and industrial views on what is missing or might be improved in the aforementioned degrees.[2]

It clearly emerged from the data collected that language and linguistics students need to develop a new skill set and a new mind frame to meet the professional challenges lying ahead. Taking the UPSKILLS needs analysis as the starting point, in this report we propose a new professional profile description, which we refer to as **language data and project specialist**. The profile contains a description of knowledge, skills and competences that present-day and future graduates in languages and linguistics should possess in order to enhance their employability. Given that there are a multitude of existing graduate profiles figures and professional roles associated with the study of language, we do not see the profile we propose here as monolithic, but rather as modular and adaptable to the needs of different study programmes and different positions in the job market. Four sub-profiles are singled out, two more oriented towards research and data analysis, and two more related to the coordination of data, people and processes; these are more specific in their typical tasks, responsibilities and competences, but nonetheless flexible.

The structure of the report is as follows. In Section 2, we sum up the results of the different components of the needs analysis. In Section 3, we focus on describing the new profile and its sub-profiles with the associated learning objectives (knowledge, skills and competences), as well as typical tasks and responsibilities. In Section 4, we provide some pointers on the educational needs that emerged from our analysis and the profile formulation. We also introduce plans for creating an interactive profiling tool that will be made available on the project website,[3] explaining how it can be used by lecturers and other stakeholders. Section 5 concludes with some final remarks on future prospects.

2. The UPSKILLS needs analysis: A summary of the different sources of information

The UPSKILLS project aims to reconcile the higher education and the market perspective on language and linguistics degrees by offering suggestions for future-proofing the language professions for the digital business sector, as well as for contemporary institutional and academic posts. Building on a small-scale study conducted during the preparation of the project

---

[2] Based on a preliminary needs analysis conducted prior to the start of the project application stage, as well as the partners' own experiences, we believe that most of the conclusions drawn also apply to the public sector and academic research. These settings, however, were not directly targeted in this needs analysis.

[3] https://upskillsproject.eu

proposal, the needs analysis that we performed captured sources of information ranging from university curricula to input from industry stakeholders. The main results of the different aspects of the analysis are reported in the following subsections.

2.1 Survey of university curricula

An important initial step in the UPSKILLS needs analysis was a survey of Bachelor and Master curricula in language and linguistics degrees implemented at European universities (Gledić et al. 2021a). This survey aimed to inform the project partners of the extent to which the skills, experience, and knowledge identified as being underrepresented in the preliminary needs analysis that preceded the formulation of the project proposal are indeed absent from the curricula of the relevant degrees and should be targeted through interventions and materials that will be designed during the project lifetime. Following the project proposal, the focus was on the presence of research skills (including the scientific method, problem solving, and project management), data acquisition skills (including data collection in experiments with human subjects, the use of linguistic corpora, and computer programming), data handling skills (including standards and repositories for data conservation, statistical analysis, and machine learning), and two cross-cutting components (linguistic theory and research management).

The survey consisted in drawing a list of European language and linguistics degrees, creating and analysing a representative sample, and performing an additional study of a selection of degrees that the partners identified as exemplary in the context of the project. The initial list of degrees was built from the didactic offer of institutions included in the QS World University Rankings in the areas of linguistics and modern languages;[4] further non-ranked degrees were added by the UPSKILLS project partners. A representative sample was drawn taking into account the country where the degree is offered, the level of study (BA vs. MA/MSc), and the QS ranking in the area of linguistics (1-50, 51-100, 101-150, 151-200, 201-250, 251-300, not ranked). The final sample included 122 degrees (primarily in the area of linguistics, but also modern languages and linguistic mediation), for which detailed information was gathered on the learning outcomes and subject lists. Twelve degrees (mostly MA/MSc) were then selected for additional study.

The analysis of the representative sample showed that practically all the skills, knowledge, and experience identified as important in the project proposal are underrepresented in the curricula descriptions, learning outcomes and subject lists; notably, no skill is mentioned in more than about one-quarter of the sample. Scientific skills and general data-related skills are the best covered ones, while the topic that was found to be the least represented (at least overtly) were data standards and repositories. In general, there were no major differences in terms of level of study, institutional ranking or country in this respect, with the exception of

---

[4] The rankings are available at https://www.topuniversities.com/university-rankings (last accessed on 13 May 2021).

programming, machine learning and linguistic theory, which were found almost exclusively in MA/MSc courses. It is also interesting to note that for some skills in particular, such as those related to research, absence may not necessarily indicate a corresponding lack of content, but rather lack of awareness that they should be explicitly mentioned.

The additionally selected degrees were found to offer a wide variety of subjects, with topics ranging from specific areas of linguistic theory to real-world practical skills. The characteristics that seem likely to have led to the perception of these degrees as exemplary – in addition to some of them being highly specialised – are flexible programmes with modular structure, and teaching methods that include a mixture of lectures, supervised individual work and hands-on experiences (in exercises, internships, etc.). Most of these degrees mention skills that are only marginally present in the representative sample analysis, such as analytical skills, data processing, programming, and machine learning. These degrees also show greater awareness of the potential for enhancing and recognising the skills learned implicitly; for instance, research skills in these programmes are mentioned in connection with activities that are not directly focused on research methods (e.g., essay writing). On the other hand, they still seem to focus primarily on academic work and traditional language professions, and do not offer substantial education in statistics and experimental design.

Overall, the study of curricula showed the need to provide additional learning content in line with the job market requirements, and to empower students and educators to identify and emphasise the skills that might be only implicitly present in existing curricula. An example of what can be done along these lines is nicely captured by the following description of the career prospects of BA graduates from the Royal Holloway University of London: "As a modern linguist, you will have excellent communication, analytical and research skills combined with the proven ability to communicate fluently, alongside practical skills such as translation and interpretation. You will have developed the kind of sensitivity to different cultures that is highly prized in the workplace."[5]

This description also refers to an aspect of the survey of curricula worth mentioning, namely future career prospects. Only 14 degree descriptions out of 122 explicitly discuss this topic. These degrees come across as designed to prepare students for a wide range of careers, including research and work in academia, but also jobs focused on data, language technologies, management, and other industries (e.g., marketing, PR, tourism), showing that some institutions do make an effort to familiarise themselves with the current job market.

---

[5] https://www.royalholloway.ac.uk/studying-here/undergraduate/languages-literatures-and-cultures/modern-languages-with-translation-studies (last accessed on 3 June 2021); see also Gledić et al. (2021a: 17).

2.2 Survey of the literature

The survey of the literature (Bernardini and Miličević Petrović 2021) focused on three types of sources that discuss the knowledge, skills and competences required of 21st-century language and linguistics students: industry surveys, institutional position papers, and academic works by scholars from languages/linguistics and education studies. The objective was to discover how the rapid changes in the job market – especially, but not exclusively, those related to the development of technology and AI – are perceived and addressed.

Recent language industry surveys were found to identify several essential competences and roles required of new graduates beyond the obvious ones (concerned with languages and cultures, linguistics, translation, terminology and so forth). The most prominent among them are related to service provision (project management, client relations, quality control), marketing and copyediting (local storytelling, creative content creation, transcreation) and – above all – technology (translation memory tools, post-editing of machine translation (MT), file format conversion, handling of mark-up languages, software and web localisation, data analysis, data cleaning, training and evaluation of MT systems, written and spoken corpora creation, automatic creation, processing and analysis of multilingual content, written and spoken language analysis and synthesis).

Institutions were found to emphasise a global societal need to focus more on transversal skills such as the ability to gather and process information critically, teamwork, digital data literacy, understanding of AI and entrepreneurship. These skills are seen as a set to be nurtured in all graduates, to guarantee that they are resilient and can adapt to rapid changes throughout their careers. Concerning language degrees in particular, it is suggested that they should cultivate cultural agility and research skills in their students. This would allow them to also target alternative career paths, thus overcoming the observed branding problem of "selling themselves" to a broad target audience of students and employers. Even translation degrees, which may not have such branding issues, are invited to revamp their priorities by targeting a growing range of language services, and familiarising their students with MT and translation into the second language as additional skills.

Scholarly reflection points to the need for technology teaching in higher education to be embedded in professional workflows that require and develop creativity, research skills and data literacy. A concrete methodological proposal in this sense is that of research-based curricula, where enquiry-based activities are devised that allow students to collaborate in teams toward a research goal. This approach leaves substantial space for tailored, self-directed learning and the possibility to make mistakes in a protected environment, preparing students for independent and less error-prone job performance later on.

Overall, the skill clusters that emerged as important for twenty-first century students of language and linguistics are:

- core disciplinary knowledge (competent use of language(s), linguistic analysis skills, translation, terminology and semantics-related competences);

- (inter)cultural awareness (awareness of cultural differences, a thorough understanding of the local context, ability to localise and personalise content accordingly);

- interpersonal and entrepreneurial skills (communication skills, teamwork, marketing skills, evaluating client and market expectations, planning skills, project management, quality control skills);

- technical skills – both basic (generic IT skills, handling text in different formats, using specialised software such as CAT tools), and advanced (process automation, familiarity with AI developments);

- data manipulation skills (ability to collect, manage, curate, clean and analyse different kinds of language data);

- research skills (critical processing of information, research design, problem solving, logical thinking, hypothetical thinking, creative and innovative thinking, evaluating technologies).

The skills from the first two clusters are typically seen as already well-covered (but see Section 2.5 below). Among the remainder, despite the common view of technology as key to modern developments, it is actually data and research skills that are perceived as opening up most opportunities, and as central for addressing the challenges posed to contemporary higher education.

2.3 Insights from job advertisements

The third component in the UPSKILLS needs analysis was a corpus-driven study of job advertisements (Ferraresi et al. 2021). This component aimed to provide an overview of the knowledge, skills and competences mentioned in job posts targeting graduates in language-related degrees or professionals with expertise in this area, as well as typical tasks and responsibilities associated with these positions. The corpus was created through manual searches and collections of job advertisements published on websites of technology companies, job announcement boards on specialised websites dedicated to linguistics, and general-purpose job platforms. Some websites were explored entirely manually, while others required focused searches using the terms "linguist", "linguistics", "data", and "language specialist". Since our project targets students and graduates in language-related disciplines, job adverts for which a degree in STEM fields was a requirement were excluded a priori.

The UPSKILLS job ad corpus features slightly fewer than 200 job advertisements that describe positions requiring a combination of language/linguistics and digital or research skills.

All texts were annotated with contextual (source of job ad, company, the original URL) and text metadata (text ID, name of the job, link to an HTML/PDF copy of the text), structural information (based on manually identified sections – job title, required qualifications, job functions), and linguistic information (part of speech and lemma). The corpus was made available via the NoSketch Engine platform of the Department of Interpreting and Translation at the University of Bologna.[6] The analysis of the corpus was carried out using Microsoft Excel and the NoSketch Engine tools, based on the annotation added to the texts. An exploratory approach was adopted, aiming to detect and classify recurrent patterns based on frequency lists extracted from different sections of the adverts. The starting point was to identify, section by section, the most frequent words, sequences of 2 to 4 words, and phrases containing a noun pre-modified by another noun or an adjective; for the most frequent phrases, collocates were generated as well.

The study of the required qualifications sections of job ads revealed three main categories of requirement: formal education, experience and knowledge, and skills and abilities. Interestingly, a university degree was not mentioned as a necessary qualification in all adverts (a Bachelor's degree was listed as a requirement in 40.6% of ads, a Master's in 35%, and a PhD in 10.7%, with overlaps). The knowledge and experience found to be most in demand are related to data, tools and techniques. They include the ability to analyse or annotate (large quantities of) language data, knowledge of specific tools or software (e.g., command-line or CAT tools), techniques to analyse or manipulate data (e.g., machine learning or regular expressions), and programming (primarily Python); these categories overlap to a large extent with the categories of data acquisition and handling skills identified as being underrepresented in the survey of curricula summarised in Section 2.1. Familiarity with linguistics (including the subfields of semantics, syntax and morphology) also features prominently, followed by knowledge of computational linguistics and NLP, and by translation and localization. Knowledge of at least one foreign language is also often required, as is experience with research activities (both quantitative and qualitative) and project management. Among the skills and abilities, in addition to those related to research, relational ones seem particularly prominent (communication and interpersonal skills). Attention to detail, organizational skills and the ability to work in a fast-paced environment and/or independently also occur frequently.

For the job functions, two broad categories emerged: linguistic, research- and technology-focused tasks, and general tasks. In the first category, perhaps somewhat unexpectedly, the most frequently mentioned item is quality assurance, which seems to be cross-cutting between linguistics, research and technology, encompassing tasks such as performing quality checks or improving the quality of NLP tools or data outputs. Linguistic tasks include translation and localization, annotation of language data, and transcription of audio files. Research tasks usually consist in collecting and analysing/categorising language data, and in research work per se (e.g. participating in the company's research and development

---

activities). The tasks that are technological in nature consist in developing NLP or other tools, testing or improving their performance, and building or training language models in the context of machine learning. Among the general tasks, the majority of adverts mention teamwork, followed by project management and customer service or support. Report writing is also rather present, often related to the ability to (analyse and) summarise data.

Finally, typical job titles were analysed, showing that – unsurprisingly – the most frequent keyword is "linguist", within the titles of "computational linguist", "(associate) linguist", and "analytical linguist" or "data linguist". "Data" was another term that featured prominently, especially with respect to "data scientist" and "data analyst" positions; "speech scientist" and "language analyst" were found as less common titles for similar positions. On the other hand, "project manager" and "language manager", as well as "project coordinator", emerged as rather frequent organisational roles open to language/linguistics graduates. It should be noted, however, that the corpus analysis did not find an entirely consistent correspondence between the job titles and the actual job profiles and task descriptions, pointing to the need for a better understanding of professional prospects and denominations.

2.4 Survey of business sectors hiring linguists and language professionals

The survey of the business sector(s) hiring linguists and language professionals was the fourth step in the needs analysis (see Gledić et al. 2021b). For this task, a questionnaire was administered to companies, with questions formulated based on the analysis of the corpus of job advertisements described above; most of the questions were closed-ended, with an "Other" option that could be used for expressing more personalised views, and with some space at the end for additional comments. The target group were employers from digital and data-intensive sectors, who do not necessarily identify language and linguistics graduates as a target talent pool of employees for their business. This group is additionally expected to benefit from the co-construction and dissemination of a specific job profile with associated competences, skills and knowledge.

A total of 70 responses were collected from companies of different sizes, coming from several European countries and from businesses ranging from language service providers to marketing and finance. Around 80% of the participating companies already had at least one position related to languages and linguistics, while 60% planned to add (more) such positions in the near future. The most represented positions were found to be those of language specialists and computational linguists or language engineers; these were followed by managers/coordinators and analytical/data linguists, and by research associates. Among less frequently mentioned occupations were product managers, recruitment officers, copy/content writers, language editors, data scientists and linguistic testers, among many others. The main tasks were identified as working with data, working with technological tools and software and

communicating with teams, clients and/or vendors, as well as conducting research and managing projects. Writing reports and evaluating tasks were also mentioned.

Crucially for UPSKILLS, the companies provided feedback on the skills they perceived as most important for their language-related positions. Among these, the most frequently selected ones were problem-solving and communication skills. Attention to detail, analytical skills and organisational skills were also indicated frequently, followed by technical skills, working under pressure, presentation skills, and creativity. Additional skills suggested by the industry were innovation, enthusiasm, willingness to learn and not be afraid to try new things, and resourcefulness. Regarding the skills perceived as needing improvement, problem-solving topped the list, followed by technical skills and organisational skills. Communication skills, attention to detail, analytical skills, working under pressure and creativity followed, with presentation skills perceived as already being at a satisfactory level. Other skills suggested by the industry were innovation, learning to work in a business environment, enthusiasm, willingness to learn, being able to adapt to new technologies, and standards. An interesting comment made by an industry representative was that "Most language experts struggle to develop strategic thinking and often focus on a small set of details" (Gledić et al. 2021b: 14), showing that attention to detail needs to be coupled with "bigger picture" skills.

When the assigned level of importance and the perceived need for improvement were compared for different skills, the highest mismatch was identified for creativity, listed as highly important by one company only, but mentioned as lacking by around one fifth of companies. The only other area where the need for improvement was judged as higher than the importance assigned were technical skills. Some domains received need-for-improvement scores commensurate with the importance assessment – organisational skills, problem-solving skills, the ability to work under pressure. Analytical skills, attention to detail, communication skills, and presentation skills were assessed as vital and not in need of much improvement. Interestingly, one company said that insight and talent analytics might now be more important in candidate selection than linguistics/computational linguistics.

As concerns the areas of knowledge and experience that companies generally look for in language and/or linguistics experts, the most important area is, unsurprisingly, knowledge of English and/or other languages. Translation and localisation come second, and computational linguistics third. Data analysis and language technology tools (including CAT tools) are also listed as essential, followed by computer science (including programming), terminology management, project management and linguistics. Additional areas of knowledge and experience suggested by the industry are machine learning and framework/policy development.

In terms of how satisfied employers are with different kinds of knowledge, language knowledge is typically seen as not lacking, and the situation is rather similar with computational linguistics and translation and localisation (which is not surprising given that they tend to be listed as core job requirements). Knowledge and experience in the domains of computer science, language technology tools, linguistics and terminology management are

seen as somewhat lacking in comparison to their importance. The only domain where a higher number of companies noted a need for improvement than the importance assigned is project management. An illustrative comment from a company reads: "Most language experts have very limited knowledge and experience processing and analysing large sets of data, and are often uncomfortable project managing, two areas that are often required when working in large organizations that work at scale" (Gledić et al. 2021b: 16).

2.5 Interviews with industry representatives

The final component of the UPSKILLS needs analysis were semi-structured interviews with eleven industry representatives (language service providers, language technology companies, but also automotive industry and insurance services), aimed at assessing whether the results obtained through the questionnaire described in Section 2.4 were indeed representative of the attitude businesses have towards language and linguistics graduates and their formation (see Assimakopoulos et al. 2021). The interviewees were shown slides with the main questionnaire findings and were asked to comment on whether these were in line with their own experience. The responses, written down in the form of structured notes, were analysed using conventional content analysis.

All interviewees agreed that there is a large demand in the industry for graduates in languages and linguistics, and that this demand is linked specifically to relevant degrees and the associated expertise, rather than to simply being speakers of specific languages. However, there is also general agreement that specialised language-related expertise should be complemented with a strong technical background. More advanced positions in particular tend to require the ability to code and/or understand how algorithms and machine learning work. In addition, language and linguistics experts with technical background are more likely to find stable long-term employment rather than freelancing and part-time opportunities typical for translators and language consultants. This notwithstanding, it should be noted that some interlocutors stated that quick learning with no disciplinary/technical background was preferred to a strong background coupled with a rigid attitude.

In terms of knowledge, skills and competences needed for the industry and identified as problematic in language degree graduates, the key item emphasised by industry representatives was problem solving, including the ability to suggest solutions based on data and to hold multiple perspectives. The problem-solving skills appear to be intrinsically interwoven with creativity, as industry representatives expect their employees to propose solutions that are not just efficient and effective, but also innovative. Analytical thinking is another related vital requirement, enabling employees to conduct thorough research, identify problems and come up with adequate solutions. As part of analytical skills, language graduates were in particular described as "not great with numbers" (A1, Assimakopoulos et al. 2021: 30). And as already mentioned above, the expectation is that technical and computational skills should be present alongside all others, not only as instrumental, but at the level of deeper understanding. It could

be said that most of these skills come together in the requirement for language graduates to be able to look at problems from "the engineering point of view" (A8, Assimakopoulos et al. 2021: 29).

On a different, but not completely unrelated note, communication and interpersonal skills were also deemed imperative, as were organisation and management skills, including quality assessment. Particular focus was placed on project management and presentation skills. Importantly, being able to present work and ideas was mentioned not only as a skill in itself, but also in relation to analytical skills, as the ability to present clearly needs to be coupled with the capacity to condense and meaningfully synthesise information; it appears that language graduates, in contrast, often fail to deliver precise presentations that have a clearly identifiable point. Sensitivity towards cultural diversity and knowing the local context and the local language with its different registers were also commented on as important. Last but not least, while the questionnaires pointed to an overall satisfactory level of disciplinary knowledge, several interviewees highlighted a lack of ability they often noted in language degree graduates to perform text quality assessment. Language graduates were also described as having a "hard time detaching from their own interests, which are often very narrow and not relevant to the task at hand" (A8, Assimakopoulos et al. 2021: 34). Rather than focusing on very specific problems of the academic type (e.g. by conducting detailed structural analyses), language graduates were constantly incited to adopt an approach in which the central element is the practical problem to be solved and the focus is on the kind of solution that is needed (which might or might not require a structural analysis).

When industry representatives were asked about the relation between higher education and workplace reality, over one half lamented that higher education is not sufficiently goal-oriented. Graduates were often described as struggling to apply their specialised linguistic or even technical knowledge to the practical tasks required for the job, and as reluctant to deal with issues outside their area of expertise. These characteristics were naturally perceived as non-desirable in settings where versatility tends to be more important than narrow specialisation. Lack of independence and an overreliance on other people's guidance were also mentioned as problematic. Overall, the view that emerged is that a solid understanding of the field is important, but that higher education curricula need to place more emphasis on the development of technical and transferable skills. This could include specialised training that will be more hands-on and will enable graduates to think outside the box and come up with their own solutions to typical industry problems. In terms of assessment, a shift from examinations towards presentations and project-based work was also proposed.

3. The *language data and project specialist* profile

The different components of the UPSKILLS needs analysis demonstrate the necessity for language and linguistics students to develop a new knowledge and skill set and a new mind frame in order to meet the societal and professional challenges lying ahead. For this to become possible, the curricula of language and linguistics degree programmes need to be adapted to include the knowledge, skills and competences that are currently underrepresented or missing, but are requested by employers, and should thus be deemed important for future-proofing the language professions for the digital business sector. As a guide toward this goal, we provide a sketch of a new professional profile that we propose to refer to as **language data and project specialist**, complete with four more specific sub-profiles – language data analyst, language data scientist, language data manager and language project manager. Even though the profile is primarily oriented towards the industry, given the general focus on big data and a growing reliance specifically on empirical language data, we expect the outlined competences to also become increasingly sought after in public institutions and academic research, making the UPSKILLS profile a tool that can be useful for planning careers beyond the digital business sector.

3.1 Method

The profiling we propose is the result of a cumulative meta-analysis of the insights gained from existing curricula, previous literature, job advertisements and opinions of industry stakeholders (expressed through questionnaires and interviews). The approach adopted was an inductive and eclectic one, mostly qualitative in nature. Such an approach was necessary because each component of the analysis used a different method and produced results in different formats, not only in terms of the qualitative/quantitative distinction, but also in terms of the categories used; for instance, the survey of the literature and the corpus analysis did not focus on searching for a predefined set of competences, while such sets were applied in the remaining three components (even though not always the same sets).

      To create the overarching profile of language data and project specialist, the main skills and roles mentioned in the individual components of the needs analysis were listed, those that overlapped to a high extent were merged, and a classification was performed based on the six clusters identified in the literature review (which were the broadest set and were judged as having an appropriate level of granularity), with the addition of transversal skills highlighted by industry representatives. Given that one of the main purposes of the profile is to guide interventions in higher education courses and curricula, a second step consisted in classifying the items associated with each cluster as knowledge, skill or competence, so as to provide a complete set of learning outcomes. An attempt was made to include as exhaustive a list as possible; however, full exhaustiveness was deemed unnecessary, and the items that were

judged as idiosyncratic or specific to single companies were left out. For completeness, typical tasks and responsibilities were also listed.

In the selection of the proposed profile name, one of the main goals was to formulate it as generally and as comprehensively as possible. The label originally proposed in the project application, "language data scientist", became too narrow for a general profile, as the needs analysis revealed a rather heavy focus on project management skills and a marked presence of coordination positions related to language tasks. We also considered the typical industry job titles identified through the corpus analysis and through questionnaires ("analytical linguist", "data linguist", "language analyst", "project manager", etc.; see Sections 2.3 and 2.4), as well as other widely used denominations such as "data scientist" and "data analyst", but found them to be either too specific or too general to capture the profile well (see also Section 3.4, where the UPSKILLS profile is compared to other related figures).[7] The term "specialist" was finally selected as sufficiently neutral, with the addition of not only "language" (as a "language specialist" would not necessarily need the technical skills highlighted by UPSKILLS), but also "data" and "project" components, arriving at "language data and project specialist".

Partly also through the choice of the profile name, the insights that emerged from the needs analysis made it evident that what was initially envisaged by the UPSKILLS partners was actually only one of several emerging profiles corresponding to the needs of the language industry. It was thus decided to add a final step in the definition of the general language data and project specialist profile and identify several sub-profiles. To this end, a split was first made based on whether the focus was more on research, or on managerial tasks, followed by a division based on the level of independence and decision-making required. The resulting sub-profiles are shown in Table 1 and discussed in more detail in Section 3.3.

		Language data analyst
Language data and project specialist	Language data researcher	
		Language data scientist
	Language data and project manager	Language data manager
		Language project manager

Table 1. The sub-profiles of the language data and projects specialist profile

3.2 Profile description

---

[7] It was also noted that the same job titles often entail different tasks and/or different levels of responsibility.

In Tables 2 and 3 below we present the detailed proposal for a new professional profile, that of language data and project specialist. In addition to its planned use in tailoring the learning content to be developed within the UPSKILLS project, the main purpose of this profile is to guide institutional and/or individual lecturers' decisions on the kind of knowledge, skills and competences to be included in language and linguistics programmes in order for their graduates to gain access to industry positions that might currently be unavailable to them, or at least perceived as unavailable. The profile is intended as a modular one, in the sense that it can be adapted to different sub-areas (linguistics, modern languages, translation and interpreting, etc.), and different study levels (Bachelor vs. Master).[8] In other words, it is meant to be used on a "pick and choose" basis (preferably through choosing some items from each domain) rather than being linked to a single educational path. It is not a proposal for an entirely new degree programme. Finally, being a sum of insights from different sources, it is also not linked to a single job title, or a single industry position.

The profile is first described in terms of learning outcomes for contemporary language-related university degrees (required knowledge, skills and competences, Table 2). This part of the profile is based on two dimensions, a vertical one focused on seven main domain clusters identified in the UPSKILLS needs analysis, and a horizontal one based on the standard constituents of learning outcomes.

DOMAIN CLUSTERS	LEARNING OUTCOMES		
	Knowledge (what the student knows/understands)	**Skills** (what the student is able to do)	**Competences** (actions the student is ready to do)
Disciplinary	- Knowledge of specific languages, including different registers - Awareness of cross-linguistic differences	- Ability to conduct linguistic analysis at different levels of language structure - Ability to work with unknown languages	- Translating/ interpreting, post-editing, localising software and Web contents - Applying theoretical knowledge to practical tasks
(Inter)cultural	- Knowledge of specific cultural contexts - Awareness of cultural differences	- Ability to understand different local contexts - Cultural agility	- Transcreating, localising and personalising content in accordance with cultural differences
Technical	- Understanding of language technology	- Ability to work with different file formats,	- Use and creation of written and spoken language

---

[8] The students themselves, as well as other stakeholders, are also most welcome to use the profile as a source of information on what is required for contemporary language-related industry jobs.

	tools (including CAT tools), machine learning, MT and AI - Understanding of methods deriving from computational linguistics or NLP - Knowledge of a programming language (preferably Python)	mark-up languages, specialised software - Ability to work on process automation, training and evaluation of automated systems - Ability to communicate with engineers	resources - Use of language technology tools - Automatic creation, processing and analysis of monolingual and multilingual content - Terminology management, use of translation memory and post-editing tools
Data-oriented	- Digital data literacy - Knowledge of statistics - Familiarity with data standards and repositories	- Ability to collect, manage, curate, clean and analyse different kinds of written and spoken language data	- Deriving conclusions from data analysis - Turning data-derived insights into decisions
Research-oriented	- Knowledge of the research process - Knowledge of research design	- Analytical skills - Logical and hypothetical thinking - Ability to review a problem, identify a solution and foresee new opportunities	- Accessing and processing information critically - Evaluating technologies
Organisational	- Understanding of entrepreneurship	- Project management skills - Planning skills - Quality control skills - Client relations skills	- Leading projects - Evaluating client and market expectations - Producing estimates - Applying quality control procedures
Transversal	- Creative and innovative thinking - Strategic thinking	- Problem-solving skills - Presentation skills - Communication and interpersonal skills - Attention to detail - Independence and quick learning	- Teamwork - Working under pressure - Report writing and presenting

Table 2. The language data and project specialist profile - learning outcomes

With regard to the disciplinary and (inter)cultural domains, it might be worth emphasising that the rather low number of items does not indicate a lack of importance, but is a consequence of the fact that the needs analysis was primarily oriented towards what is lacking (in the existing curricula, as well as in the view of industry representatives). In addition, due to

the industry focus of the analysis, some sub-disciplines are not captured at all (e.g., language pedagogy), but this does not remove the degrees in these sub-disciplines from the addressee list of the profile. The readers are also reminded that the "familiarity with data standards and repositories" item was not directly derived from the industry requirements, but suggested as an important area by the project partners beforehand, and subsequently identified as entirely missing from current higher education curricula. By including this item in the language data and project specialist profile, the UPSKILLS consortium confirms their expectation that data standards and the way data are conserved are on their way to becoming crucial for digital businesses, as they already are for public institutions and academic research (see e.g. the forthcoming UPSKILLS report by Miličević Petrović, Gledić and Đukanović, on available teaching/learning resources, where data standards and repositories are represented through multiple projects and coordinated European initiatives, as well as the ELIS 2021 report,[9] where an increase in language asset management needs is noted).

In Table 3, the learning outcomes are associated with the typical tasks and responsibilities that graduates of language-related degrees are entrusted with in the language industry. The table shows the domains of knowledge, skills and competences in the order of importance for different tasks. A selective list is reported to avoid repetition, as a high level of disciplinary, (inter)cultural and transversal skills is vital as a basis for all tasks.

Typical tasks and responsibilities	Core knowledge, skills and competences required
Linguistic data collection	Data-oriented, technical, research-oriented, organisational
Transcription of audio files	Data-oriented, technical, research-oriented, organisational
Linguistic annotation	Data-oriented, technical, research-oriented, organisational
Translation, interpreting, localisation, post-editing	Technical, data-oriented, research-oriented, organisational
Exploratory work focusing on language data	Data-oriented, research-oriented, technical, organisational
Language data analysis	Data-oriented, research-oriented, technical, organisational
Language data research	Research-oriented, data-oriented, technical, organisational
Research on business processes and market needs	Research-oriented, data-oriented, organisational, technical

---

[9] https://euatc.org/news/full-slide-deck-for-european-language-industry-survey-available/

Work with software and technological tools (development, analysis, testing)	Technical, data-oriented, research-oriented, organisational
Work with machine learning models (development, testing and improvement)	Technical, data-oriented, research-oriented, organisational
Communication with teams, clients and/or vendors	Organisational, data-oriented, technical, research-oriented
Project management	Organisational, data-oriented, technical, research-oriented
Processes evaluation	Organisational, data-oriented, technical, research-oriented

Table 3. The language data and project specialist profile - tasks and responsibilities

3.3 Sub-profile descriptions

The profile of language data and project specialist outlined in Section 3.2 is deliberately phrased in very general terms, so that it can capture the wide range of knowledge, skills and competences identified as important for contemporary jobs in the digital business sector. However, skill and position clustering, along two lines, were also noted in the job advertisements and in feedback received from the industry representatives. Firstly, while disciplinary, (inter)cultural, technical and transversal skills were listed as equally required for all positions, some positions were found to entail more research and data skills, and some more organisational skills (in particular project management). Secondly, a distinction can be noted between more junior positions that mostly involve performing very specific and partly isolated tasks, and more senior positions that entail more independence, more responsibility and more of a "bigger picture" approach.

To capture this clustering, we propose four different sub-profiles, shown in Table 1 above. Keeping in mind the distinction between research-oriented and organisational roles, we distinguish between mid-level profiles of **language data researcher** and **language data and project manager**. Adding to this dimension the distinction in seniority, we single out two types of researchers – the language data analyst and the language data scientist, and two types of managers – the language data manager and the language project manager. These sub-profiles are described in turn in the following paragraphs.[10]

The **language data analyst** typically works on operative tasks such as language data collection, transcription and annotation, language data exploration and analysis, often in connection with translation-related tasks. Within these tasks, (s)he can be involved in

---

[10] Note that the sub-profiles aim to capture a wider range of tasks and competences than those associated with individual positions sometimes advertised under the same labels. It also goes without saying that specific positions and tasks can differ based on company size, main business and other factors.

discovering patterns and trends in language data (often using statistical analysis), and is likely to be expected to communicate these insights through reports and other types of data summarisation and visualisation.[11] The language data analyst possesses a wide range of data-oriented skills, as well as some research skills and at least a basic set of organisational skills. If (s)he has the right technical skills, involvement in the testing of language technology and machine learning models is common. The positions (s)he occupies tend to be open to BA graduates, and over time can lead to advancements to senior positions linked to the language data scientist profile.

The **language data scientist** normally conducts research on language data, and/or business processes and market needs. (S)he is mostly involved with problem solving, through asking questions, obtaining the relevant data, exploring the data, modelling the data (statistically) and communicating the results.[12] This sub-profile corresponds most closely to the profile initially envisaged by the UPSKILLS partners, and it is possibly the most attractive one in terms of opportunities, being higher up in the corporate ladder and possibly allowing for a lot of personal initiative and creativity. It requires advanced data-oriented and research-oriented skills, jointly with a relatively high level of organisational skills. It becomes particularly promising in the presence of advanced technical knowledge, skills and competences, allowing for technological tools and machine learning to be not only used, but also developed. The positions occupied by the language data scientist are mostly open to MA/MSc graduates, or individuals with sufficient experience in the field; however, in the case of adequate and timely preparation, they could also become available to BA graduates.

The **language data manager** also works primarily with data. However, her/his main focus is on data cleaning, curation and management (e.g., as part of database or translation memory maintenance, or as a follow-up on transcription and annotation tasks). The requirements comprise substantial data-oriented and organisational skills, with particular relevance given to digital data literacy and familiarity with data standards and repositories on the one hand, and planning and quality control on the other. As a junior profile, the language data manager is primarily associated to a BA degree as a formal qualification.

The **language project manager** profile emerged from the insights gathered in the needs analysis, where managerial roles and tasks were frequently mentioned. Language project managers coordinate different types of company projects and workflows, from translation tasks to the development of new language technologies. Their responsibilities also include process evaluation and communication with teams, clients and/or vendors, and they differ from the remaining sub-profiles in being more directed towards processes and people than towards data. Language data managers primarily need very advanced organisational skills, coupled with a certain level of research-oriented and data-oriented skills, as familiarity with what goes on within a company's data and research activities is necessary for being able to manage projects.

---

[11] See also e.g. https://medium.com/@udemeudofia01/disciplines-in-data-science-a1da93306528
[12] See also e.g. https://medium.com/dataseries/what-does-a-data-scientist-do-a6553dc720f

As with the language data scientist, this sub-profile is mostly for MA/MSc graduates, or for individuals with previous industry experience as language data managers or related roles.

3.4 Comparison with similar profiles

To further underscore the novelty of the UPSKILLS profile, we compare it to three related already established profiles. The first one is "computational linguist". Computational linguistics was brought up multiple times by the industry representatives in questionnaires and interviews, in reference to either positions in the company, or knowledge. While we do consider its basic understanding to be relevant for most of the contemporary language degrees, we do not equate the UPSKILLS profile, which aims to maintain the linguistic focus, with one that is almost entirely oriented towards developing computational models for the automatic processing of language data (often without much linguistic information).

The second relevant profile is that of the "digital linguist", defined in detail within the DigiLing project.[13] This profile, associated to the emerging field of digital linguistics, entails competence in at least two languages, mediation-related competences, an understanding of language analysis, an understanding of NLP techniques, basic programming skills, as well as digital content authoring and an understanding of related ethical and legal issues (Vintar et al. 2017). In addition to being more specific than the UPSKILLS profile, the digital linguist profile is connected to a dedicated degree programme at the MA level. While overlaps do exist between the learning outcomes defined by this programme (*ibid*: 19) and the UPSKILLS profile, we see the flexibility and the applicability to numerous different programmes, including those at the BA level, as major innovations of the language data and project specialist.

Finally, we believe that the UPSKILLS (sub-)profile(s) cannot be merged with the general "data scientist", "data analyst" or "project manager" profiles, as these do not include the special kind of disciplinary knowledge that is needed for handling language data. While language data are not the only data with an internal structure, they require highly specialised analyses that can only be conducted by professionals with a strong background in languages and/or linguistics. In other words, while language and linguistics students can certainly choose to take up appropriate training to specialise in computational linguistics, digital linguistics or data science, this is not necessarily contained within the profile of language data and project specialist, but can rather be perceived as a set of possible foci that the profile prepares for.

---

[13] http://www.digiling.eu/

4. How to use the *language data and project specialist* profile

The two main uses planned for the profile and the sub-profiles outlined in Section 3 concern the education of students enrolled in language-related degrees. The first intended use is internal to the UPSKILLS project, where the final selection of topics for the new teaching and learning content to be developed (or adapted from already existing contents) will be largely based on the profile(s). The second one goes beyond UPSKILLS and concerns the use of the profile by stakeholders involved with language and linguistics programmes at different universities (lecturers, institutional decision-makers and even students themselves), who will be invited to utilize the profile for detecting the knowledge, skills and competences that could be useful for their students' career prospects but are not yet taught; since the profile will be directly connected to the learning content produced by UPSKILLS, the invitation will be extended to integrating this content in the existing curricula. An additional step will be the promotion of the profile with the industry, whose representatives helped co-construct it. Bringing the industry closer and making them more aware of the developments in the higher education sector could lead, on the one hand, to a more effective implementation of the profile by universities, and on the other hand to a revision of industry views on where their main target pool of employees lies, with enhanced employability of language graduates as a joint outcome.

4.1 Next steps for UPSKILLS: from needs analysis to new learning content

In relation to new learning content, the UPSKILLS project planned for the creation of materials in three main domains: research skills (decomposed in research methods, problem solving, and project management), data acquisition skills (divided in collecting data from human subjects, text/speech processing, and programming), and data handling skills (composed of language data analysis, machine learning, and data standards and repositories). Despite the different groupings, these domains match a substantial part of the newly proposed profile of the language data and project specialist. In addition, they were confirmed to be underrepresented in the relevant higher education degree programmes, some being absent altogether, others present in a low number of degrees, and others yet limited to the MA/MSc level.

Judging from the needs analysis, the emphasis in learning content creation should be placed primarily on data analytics, research skills and project management, where the latter is not to be seen as a side requirement for conducting research efficiently, but as a skill highly valued in itself that can be the foundation for a successful non-research career in the digital business sector. Data analytics requirements appear to be largely quantitative in nature, confirming the need to introduce (more) training in statistics in language-related degrees, without forgetting, however, that qualitative analyses and general analytical skills are no less central. Technical skill, expectedly, also surfaced among the core requirements. The fact that multiple industry representatives expressed the view that language graduates should be able

not only to use technology, but also to understand how it works, corroborates the UPSKILLS proposal to introduce programming and machine learning already at BA level.

One particular element that was also judged as essential, and that we believe should be mentioned separately, are problem-solving skills. Problem solving was very often evoked in the needs analysis as both indispensable and lacking. And while it was included in the USPKILLS plans from the beginning, similarly to project management, it turned out to be much more than a support component to research skills, as it is required for practically every task mentioned. The overarching impression thus seems to be that problem solving should be dealt with not only in a separate component of the learning content, but also within other components; it can certainly tie in well with the data-oriented and research-oriented skills, but also with project management. It would in addition be desirable to think of ways to better connect problem-solving skills with disciplinary knowledge, as multiple industry representatives remarked that language graduates do not lack disciplinary background, but very much lack the ability to apply their knowledge to practical tasks. The framework of research-based teaching and learning, which is also highlighted as appropriate in the needs analysis and originally planned as part of the work on the UPSKILLS learning content, could provide a suitable setting for that.

Finally, two special cases deserve to be mentioned. First, the topic of data standards and repositories was found to be completely absent from current curricula and was not brought up by industry representatives. Second, the industry did not mention data collection from human subjects either, and this topic was also found to be fairly rarely included in curricula. Given the industry focus of the needs analysis, neither case is unexpected, as data standards are more present in public institutions, and experiments tend to be associated primarily with academic research. However, as already mentioned in section 3.2, we believe that data standards will become more relevant for the industry as well (they currently appear to be mostly required under general IT skills, in reference to file formats), and we thus propose to keep them both in the language data and project specialist profile and in the learning content.[14] Skills related to experimental studies with human subjects, on the other hand, could be helpful in organising annotation campaigns and data collection processes; note, however, that we do not have empirical evidence in our data analysis to confirm this. Another possible advantage of keeping data collection from human subjects as a topic in the UPSKILLS learning content is that it could provide a fertile ground for research-based teaching and for structured tasks in which theoretical knowledge must be applied to real-world problems. We leave the final decision on this for the next steps in the project, in particular intellectual outputs 2 and 3, which investigate best practices, compile dedicate guidelines for research-based teaching and define the learning content to be created.

---

[14] Judging from the survey of curricula, including data standards and repositories in the UPSKILLS learning content can be seen an asset for the project, as well as an opportunity to promote this topic with industry.

4.2 Addressing educational needs beyond UPSKILLS

As concerns the use of the language data and project specialist profile beyond the UPSKILLS project, the main target audience consists of higher education stakeholders in the broad areas of languages and linguistics. Representatives of institutions, lecturers, and other interested parties can use the profile as a source of information on the current expectations of the digital business sector that employs language and linguistics graduates, and consequently as a guide towards curricula enrichment. In other words, the profile does not require entirely new degree programmes or new curricula; it is intended to help in formulating courses or modules that can be added to already existing curricula in order to complement them.

In using the language data and project specialist profile, it is important to always keep in mind that it is an abstraction based on different job profiles encountered throughout the needs analysis, i.e., that it is a fluid framework rather than a definitive solution. When used for planning new courses or selecting new topics to add to old courses, the profile should be gauged against the contents and objectives of specific degree programmes and/or individual subjects; not all learning objectives will be relevant for every degree programme, or for all the specific professional profiles they prepare their students for. It is thus not possible, or in general deemed necessary, to add all components to every programme. The four sub-profiles can be a useful tool for a more focused selection.

Importantly, the role of UPSKILLS does not end with the profile(s) description. During the project lifetime, the partners will work on reusing already existing open materials, and on creating and making available new learning content that addresses the needs identified through the profile(s), so that colleagues from the partner institutions and from other universities will have the possibility to use and adapt a large set of open materials, some of which game-based, instead of necessarily creating their own from scratch. As part of dissemination activities and in the intellectual output dedicated to the learning content, an interactive profiler will also be created and published on the project website, which will ask as input information on the type of degree programme, on the target student careers, and on the knowledge, skills and competences that are already included in the curricula. The output will be information on what might be missing based on the language data and project specialist profile and its sub-profiles; this information will be linked to any relevant UPSKILLS learning content that could be helpful in managing the missing elements.

In addition, the higher education stakeholders will have access to the UPSKILLS best practices in research-based teaching, in which enquiry-based and problem-solving activities are devised for students to collaborate in teams toward a research goal. This teaching method emerged as an often-suggested paradigm for tackling the missing skills (especially those related to problem solving), and for complementing the skills that are already present in the curricula but would benefit from an "uplift" (e.g., applying theoretical knowledge to real-world practical tasks). The learning content and the best practices will also tackle the issue of how to better

emphasise the knowledge and skills that are implicitly present in the curricula, but are not explicitly stated through course titles or learning outcomes.

Lastly, even though the suggested sub-profiles are stratified, and some appear to be more relevant for MA/MSc graduates, the UPSKILLS consortium will continue to promote a focus on the BA level and work on moving at least one part of the contents and outcomes that tend to be present from the MA/MSc onward (if taught at all) to an earlier level of study.

5. Concluding remarks

The UPSKILLS needs analysis conclusively showed that students in language-related fields need to develop the ability to adapt, diversify their skills, be flexible and resilient. Despite the variety of data sources consulted, as well as differences between the companies that took part (in terms of their specific activities, size and requirements for employees), an overarching profile labelled "language data and project specialist" emerged, as a collection of typical knowledge, skills, competences, and tasks. This profile describes a linguistics or language graduate that remains clearly distinguishable by the core disciplinary and (inter)cultural competences, but who is able to use this knowledge in different practical tasks and who is at the same time (at least to some extent) familiar with contemporary developments in technology and AI, skilled in language data collection, has knowledge of quantitative and qualitative data analysis, excels at problem solving, and possesses organisational skills such as project management. Four sub-profiles were also defined, based on predominance of research-oriented vs. organisational skills and the degree of independence and control (language data analyst, language data scientist, language data manager, language project manager).

In the next step, the findings from the profile and from the needs analysis in general will be used to decide on the final selection of learning contents to be adapted or created anew and shared with the higher education community. The higher education stakeholders will be encouraged to work on providing opportunities for their students to acquire this new skill set, by including in their teaching those contents and learning outcomes that are currently missing, and by promoting research-based activities as much as possible. Lecturers can also help through raising awareness about the job market options that could be perceived as irrelevant by many students, and they could act as motors for change and mediators between the often traditional and slow to adapt university programmes and the rapidly evolving market requirements, in which disciplinary knowledge is just a basis to be complemented with a series of more general competences.

References

Assimakopoulos S., Vella M., van der Plas L., Milicevic Petrovic M., Samardžić T., van der Lek I., Pallottino M., Bernardini S., Ferraresi A. (2021). Graduate skills and employability: Focus interviews with selected job market stakeholders. *UPSKILLS Task Report*. https://doi.org/10.5281/zenodo.5030913

Bernardini S., Miličević Petrović M. (2021). Toward a new profile for twenty-first century language specialists: Industry, institutional and academic insights. *UPSKILLS Task Report*. https://doi.org/10.5281/zenodo.5030873

Ferraresi A., Aragrande G., Barrón-Cedeño A., Bernardini S., Miličević Petrović M. (2021). Competences, skills and tasks in today's jobs for linguists: Evidence from a corpus of job advertisements. *UPSKILLS Task Report*. https://doi.org/10.5281/zenodo.5030879

Gledić J., Miličević Petrović M., van der Lek I., Bernardini S., Đukanović M., Assimakopoulos S. (2021). Survey of curricula: Linguistics and language-related degrees in Europe. *UPSKILLS Task Report*. https://doi.org/10.5281/zenodo.5030861

Gledić J., Budimirović J., Đukanović M., Samardžić T., Jukić S., Ferraresi A., Bernardini S., Aragrande G., van der Plas L., van der Lek I., Soldatić N. (2021). Survey of business sectors hiring linguists and language professionals. *UPSKILLS Task Report*. https://doi.org/10.5281/zenodo.5030891

Miličević Petrović M., Gledić J., Đukanović M. (under review) Overview of existing materials. *UPSKILLS Task Report*.

Vintar Š., Matuška O., Škerl J., Bago P., Ciobanu D., Cvrček V., Čulo O., Jakubiček M., Mikelić Preradović N., Nitzke J., Secara A., Žabokrtsky Z. (2017) Labour market needs survey and the DigiLing model curriculum. *DigiLing Project Report*. http://www.digiling.eu/wp-content/uploads/2017/05/DigiLingReport_IO1.pdf