



Co-funded by the  
Erasmus+ Programme  
of the European Union



# IO1: “Language data and project specialist”

From an UPSKILLS profile to  
classrooms across Europe

Adriano Ferraresi & Silvia Bernardini, UNIBO

# A brief recap on IO1 – Needs analysis

UPSKILLS 1st Intellectual Output aimed to explore

- the **current academic offer** in language-, linguistics- and translation-related fields
- the **requirements** that the **job market** has for graduates in these areas

# How?



## **Review of the literature**

- Academic works
- Language industry surveys
- Institutional reports / position papers

# How?



## Survey of existing curricula

- 122 BA & MA degrees in languages and linguistics



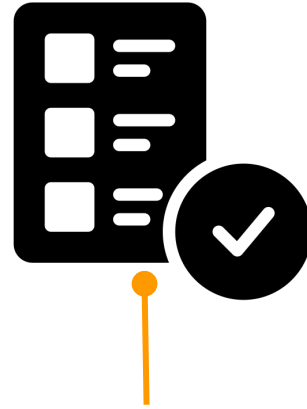
# How?



## Corpus-based analysis of job advertisements

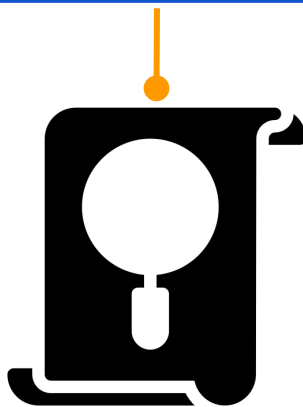
- ~ 200 job ads from 112 companies

# How?

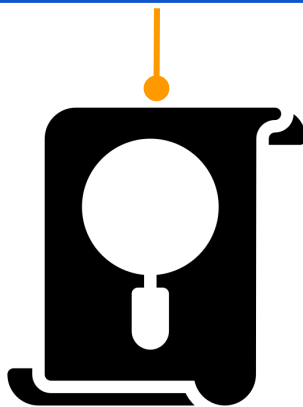
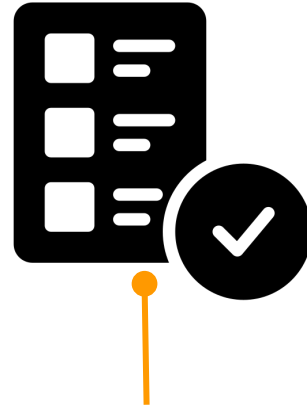


## Online questionnaire

- 70 businesses, from language service providers to marketing and finance companies



# How?



## Focus groups

- 12 job market stakeholders from 11 different companies

# Some insights

- **Research** skills, **data acquisition** skills and **data handling** skills are **underrepresented in curricula descriptions**
- Global **societal need** to focus more on **technological** and **transversal skills** (e.g. critical thinking, digital data literacy)
- **Problem-solving, communication skills** and **independence** are the most needed (and in need of improvement) skills according to companies
- There is a **large demand** in the industry for graduates in languages and linguistics: twists to the “old” occupations, but also entirely new options



# *Language data and project specialist*

A new **modular profile** for graduates in language-related disciplines, summarising:

- required knowledge, skills and competences
- typical tasks associated with emerging jobs

Split into 4 main sub-profiles



## Language data analyst

Language data collection,  
annotation, analysis



## Language data scientist

Language data and processes  
research



## Language data manager

Language data cleaning,  
curation, management



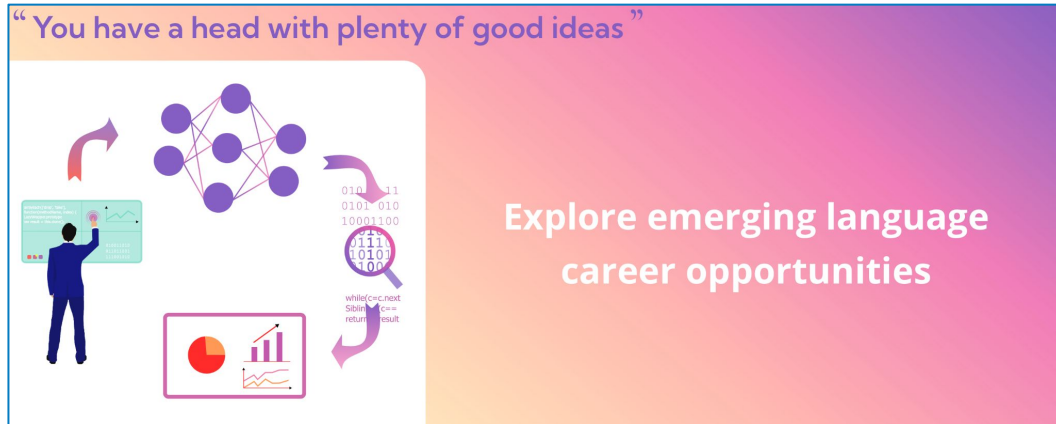
## Language project manager

Language project and  
workflow coordination

# The full IO1 report



# The interactive version of the profile



<https://upskills-dev.dipintra.it>

A website intended for **students** AND **lecturers**, including:

- list and description of required knowledge, skills and competences
- description of the professional (sub-)profiles and associated tasks
- survey

Required knowledge, skills and competences

# WHICH KNOWLEDGE, SKILLS AND COMPETENCES?

01 DISCIPLINARY

Knowledge of languages, linguistic analysis, translating and interpreting

[Read more](#)

02 (INTER)CULTURAL

Awareness of cultural differences, understanding local contexts, mediating between them

[Read more](#)

03 TECHNICAL

Using and developing language technologies together with language engineers

[Read more](#)

04 DATA-ORIENTED

Familiarity with language data and their collection, curation and (statistical) analysis

[Read more](#)

05 RESEARCH-ORIENTED

Applying analytical and critical skills to approach problems and advance knowledge

[Read more](#)

06 ORGANISATIONAL

Planning, managing and leading projects and teams

[Read more](#)

07 TRANSVERSAL

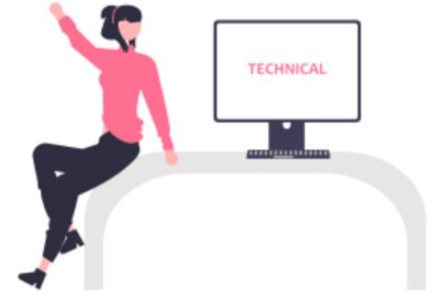
Thinking and communicating creatively and efficiently within teams and with clients

[Read more](#)

## Technical

### Using and developing language technologies together with language engineers

This means that you know about different types of language technologies, such as CAT tools, machine translation engines and ChatGPT, and you know not only how to use them, but also have some knowledge of how they work; you also have some programming skills, e.g. in Python



## Data-oriented



### Familiarity with language data and their collection, curation and (statistical) analysis

This means that you have some experience in collecting linguistic data (e.g. from corpora or from speakers, in an interview or an experiment), in curating them and storing them following a set of standards, and in analysing them (especially quantitatively)

# The professional (sub-)profiles

## PROFILE DESCRIPTION

As a **language data analyst** you might work on operative tasks such as language data collection, transcription and annotation, language data exploration and analysis, often in connection with translation-related tasks. You probably possess a wide range of data oriented skills, as well as some research skills.

### TYPICAL TASKS AND RESPONSIBILITIES

- ✓ Linguistic data collection
- ✓ Transcription of audio files
- ✓ Linguistic annotation
- ✓ Language data analysis
- ✓ Exploratory work focusing on language data

### KNOWLEDGE, SKILLS AND COMPETENCES

Organizational Collect Bachelor's  
**Statistical** Reporting  
Data exploration Transcription  
Test **Texts** Annotation

#### #1 Data-oriented

**Familiarity with language data and their collection, curation and (statistical) analysis**

This means that you have some experience in collecting linguistic data (e.g. from corpora or from speakers, in an interview or an experiment), in curating them and storing them following a set of standards, and in analysing them (especially quantitatively)





## PROFILE DESCRIPTION

As a **language project manager** you coordinate different types of company projects and workflows, from translation tasks to the development of new language technologies. Your responsibilities also include process evaluation and communication with teams, clients and/or vendors. You are directed more towards processes and people than towards data.

### TYPICAL TASKS AND RESPONSIBILITIES

- ✓ Communications with teams, clients and vendors
- ✓ Project management
- ✓ Processes evaluation

In the following sections, you will find more details about what you need to know to effectively perform previously presented tasks. Please have in mind that maintaining a strong foundation of [disciplinary](#), [\(inter\)cultural](#) and [transversal](#) skills is equally indispensable for success across all listed tasks and responsibilities.

### KNOWLEDGE, SKILLS AND COMPETENCES

Master's Processes  
**Workflows** Data  
Evaluation Innovation  
**Organizational** Research  
Problem-solving **People skills**

#### #1 Organisational

##### Planning, managing and leading projects and teams

This means that you have some familiarity with entrepreneurship, you are able to manage different kinds of projects, plan their implementation, implement quality assurance procedures and ensure that teams work well together towards a common goal



## PROFILE DESCRIPTION

As a **language data scientist** you normally conduct research on language data, and/or business processes and market needs. You are mostly involved with problem solving, through asking questions, obtaining the relevant data, exploring the data, statistically modelling the data and communicating the results.

### TYPICAL TASKS AND RESPONSIBILITIES

- ✓ Language data research
- ✓ Research on business processes and market needs
- ✓ Work with machine learning models
- ✓ Work with software and technologies tools

### KNOWLEDGE, SKILLS AND COMPETENCES

Organizational Communicate  
Innovation Design  
Statistical Modeling  
Develop Master's Management  
Research

#### #1 Data-oriented

Familiarity with language data and their collection, curation and (statistical) analysis

This means that you have some experience in collecting linguistic data (e.g. from corpora or from speakers, in an interview or an experiment), in curating them and storing them following a set of standards, and in analysing them (especially quantitatively)



## PROFILE DESCRIPTION

As a **language data manager** you work primarily with data: data cleaning, curation and management (e.g., as part of database or translation memory maintenance). You need to be familiar with data standards and repositories on the one hand, and planning and quality control on the other.

## TYPICAL TASKS AND RESPONSIBILITIES

- ✓ Work with software and technologies tools
- ✓ Transcription of audio files
- ✓ Translation, interpreting, localisation, post-editing

## KNOWLEDGE, SKILLS AND COMPETENCES



### #1 Technical

Using and developing language technologies together with language engineers

This means that you know about different types of language technologies, such as CAT tools, machine translation engines and ChatGPT, and you know not only how to use them, but also have some knowledge of how they work; you also have some programming skills, e.g. in Python

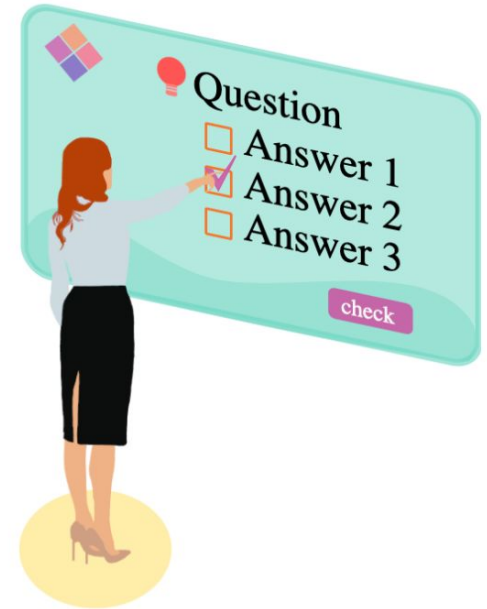


# The survey

## WHICH PROFILE SUITS YOU BEST?

Your studies focus(ed) on languages, linguistics or translation/interpreting.  
You have solid knowledge of cultural issues attached to specific languages.  
You have good communication skills. It seems that you might fit the professional profile of *Language Data and Project Specialist!*  
Do our short survey and learn more about the knowledge and skills you already have, those that might be useful to acquire, and discover a more specific subprofile that describes you.  
This can also help you understand the kinds of jobs that you can apply for and the kinds of tasks they involve.

[Get started](#)



The core disciplines in my degree are/were in the area of:

- a Linguistics: general, applied, comparative...
- b Languages, cultures, literatures
- c Translation, interpreting, intercultural communication
- d Not related to languages

In my future job, I might enjoy:

- a Reflecting on how language works
- b Looking for patterns in language data

I think I am able to:

- a Ask questions about language
- b Examine data from texts or speaker
- c Handle projects related to language
- d Collect and catalogue language data

I am especially eager to learn more about:

- a Linguistic research
- b Text and speech processing
- c Statistical analysis
- d Artificial intelligence



## Language data manager

As a **language data manager** you work primarily with data: data cleaning, curation and management (e.g., as part of database or translation memory maintenance). You need to be familiar with data standards and repositories on the one hand, and planning and quality control on the other.

[Continue reading](#)

Please note that this profile name may differ from job profiles in the labour market.

In addition to the mentioned skills, for your future career it will be important to continuously enhance your [disciplinary](#), [\(inter\)cultural](#) and [transversal](#) skills.

# Survey results (suggested sub-profile)

## UPSKILLS Moodle courses

Take a look at these UPSKILLS Moodle courses. They can help you deepen your knowledge in the areas relevant to you.

### Upskilling your "Introduction to Language variation" course



The material presented in this block amounts to a full introduction to language typology/language variation course. The workload associated with the entire block...

[Read on Moodle](#)

### Collecting language data from human participants



This 6 ECTS learning block is dedicated to collecting language data from human participants in the areas of morphophonology/morphosyntax...

[Read on Moodle](#)

# RESULTS FROM SURVEY PARTICIPANTS

LANGUAGE DATA ANALYST



LANGUAGE DATA SCIENTIST



LANGUAGE DATA MANAGER



LANGUAGE PROJECT MANAGER





## MEET THE TEAM



**Silvia Bernardini**

*Full Professor*

[Go to personal web site](#)



**Maja Miličević Petrović**

*Associate Professor*

[Go to personal web site](#)



**Adriano Ferraresi**

*Associate Professor*

[Go to personal web site](#)



**Olga Arsic**

*Research fellow*

[Go to personal web site](#)



**Daniele Negretti**

*Web developer*



Co-funded by the  
Erasmus+ Programme  
of the European Union



Thank you