



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



# Dissemination plan and materials for the UPSKILLS learning content

*UPSKILLS Intellectual output 3.4*

Compiled by:

Iulianna van der Lek\*, Tihana Kraš†, Stavros Assimakopoulos‡ and Jelena  
Gledić\*\*

\*CLARIN ERIC

† University of Rijeka

‡ University of Malta

\*\*University of Belgrade

**UPSKILLS: UPgrading the SKILLS of Linguistics and Language Students**

Erasmus+ Programme

Key Action 2: Cooperation for Innovation and the Exchange of Good Practices Action

KA203: Strategic Partnerships for Higher Education



Grant Agreement Number: 2020-1-MT01-KA203-074246

## UPSKILLS Consortium:



&

with financial support from



Dissemination plan and materials for the UPSKILLS learning content © 2023 by the authors (UPSKILLS Consortium) is licensed under Attribution-ShareAlike 4.0 International. To view a copy of this license, visit <http://creativecommons.org/licenses/by-sa/4.0/>



### **Disclaimer:**

The European Commission's support for the production of this publication does not constitute an endorsement of the contents, which reflect the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

## Table of contents

Table of contents	3
List of figures	4
Executive Summary	5
<b>1. Introduction</b>	<b>6</b>
<b>2. Dissemination within the partner institutions</b>	<b>6</b>
2.1 University of Belgrade	6
2.2 University of Bologna	7
2.3 CLARIN ERIC	8
2.4 University of Geneva	9
2.5 University of Graz	9
2.6 University of Malta	10
2.7 University of Rijeka	10
<b>3. Dissemination outside the Partner Institutions</b>	<b>12</b>
3.1 University of Vienna, Austria	12
3.2 KU Leuven University, Belgium	12
3.3 University of Osijek, Croatia	12
3.4 University of Zagreb, Croatia	13
3.5 Vytautas Magnus University, Lithuania	13
3.6 University of Ljubljana, Slovenia	14
<b>4. Dissemination via Open Registries and Repositories</b>	<b>14</b>
4.1 The Digital Humanities Course Registry	15
4.2 The SSH Open Marketplace	17
4.3 DARIAH Campus	19
4.4 OER Commons	20
4.5 Zenodo	22
4.6 CLARIN.SI Repository	23
<b>5. Dissemination plan and materials</b>	<b>25</b>
<b>Conclusions</b>	<b>25</b>
<b>References</b>	<b>27</b>
Annex A: UPSKILLS dissemination e-mail template	28
Annex B: UPSKILLS conclusion press release	29
Annex C: UPSKILLS description slide deck	31
Annex D: UPSKILLS brochure	36
Annex E: UPSKILLS learning content flyers	37

## **List of figures**

<i>Figure 1. Browsing interface of the Digital Humanities Course Registry</i>	14
<i>Figure 2. The Search Interface of the SSH Open Marketplace</i>	16
<i>Figure 3. How to Access the SSH Open Marketplace</i>	17
<i>Figure 4. The Search Interface of DARIAH-CAMPUS</i>	18

## **Executive Summary**

This document outlines the plan for sharing the learning content developed in Task 3.2 of UPSKILLS Intellectual Output 3. According to the data collected via three surveys, the learning content will be integrated into the curricula of **seven (out of eight) partner institutions** (including partners funded by Movetia) and **six universities** outside the UPSKILLS consortium. Additionally, we include a brief analysis of existing educational registries and repositories, which are suitable for further disseminating the UPSKILLS learning content. Finally, we have created promotional materials such as a press release, digital brochures, flyers and videos to increase awareness of available resources.

## 1. Introduction

Task 3.4 involves disseminating the UPSKILLS learning content within and beyond the partner institutions, where one of the main activities is including the developed material in relevant course registries and repositories. The components of the task are the following:

- identifying partner institutions (funded by EU and Movetia) that will integrate the UPSKILLS learning content in their university curricula;
- identifying institutions beyond the consortium interested in integrating the UPSKILLS learning content into their university curricula;
- identifying registries and repositories for open educational resources, including their requirements, which would be suitable to include the UPSKILLS learning content;
- the inclusion of the UPSKILLS learning content in selected registries and repositories for open educational resources;
- the creation of dissemination materials (a press release, an email template, a registration form for using the UPSKILLS learning blocks, a PPT presentation, a press release, digital brochures, flyers and promotional videos).

## 2. Dissemination within the partner institutions

A survey was created and sent out to the project partners to find out who is interested in integrating the UPSKILLS learning content into their university curricula. **Seven partners** have expressed their interest, as outlined below. In some cases, project members' intention to use the UPSKILLS learning content is based on prior piloting; in others, it is not. In this section, we report on this piloting where it took place, and on the UPSKILLS partners' plans to include our materials in their university curricula.

### 2.1 University of Belgrade

Starting from the school year 2023-2024, project members **Maja Đukanović, Jelena Budimirović** and **Jelena Gledić** will use the following learning blocks:

- *A Glimpse into Language Data Science*
- *Analytical Thinking and Problem-Solving*
- *Collecting Language Data from Human Participants*
- *First Steps into Scientific Research*
- *Introduction to Language Data: Standards and Repositories*

- *Processing Texts and Corpora*
- *Project Management*
- *The Essence of Machine Learning for Linguists in Tech*

The blocks will be used in their teaching in BA courses at the Faculty of Philology by incorporating modules of the learning content into existing courses and offering them for guided self-study to students who need to catch up or wish to learn more. Upon successfully applying this approach, the materials will be promoted to other lecturers at the Faculty of Philology and other schools of the University of Belgrade.

## 2.2 University of Bologna

About 10 to 15 lecturers in the Department of Interpreting and Translation are interested in using one or more of the following learning blocks:

- *A Glimpse into Language Data Science*
- *First Steps into Scientific Research*
- *Collecting Language Data from Human Participants*
- *Introduction to Language Data: Standards and Repositories*
- *Processing Texts and Corpora*
- *Start Programming with Python in 10 Steps*
- *The Essence of Machine Learning for Linguists in Tech*

Project members **Silvia Bernardini, Adriano Ferraresi, Alberto Barrón-Cedeño and Maja Miličević Petrović** will use parts of these learning blocks to teach optional 6 ECTS courses *Introduction to Computational Thinking* (1st year), *Text Processing* (2nd year) and *Researching Language* (3rd year) within a reformed BA degree course *Languages and Technologies for Intercultural Communication* and an UPSKILLS-based BA degree course *Language, Technology, Research*, as well as to teach 5-6 ECTS courses *Corpus Linguistics*, *Language Data Analysis*, *Natural Language Processing*, and *Profession-Based Research* within the *Translation and Technology* component of the MA degree course *Specialized Translation*. They will also try to implement parts of the learning content into existing BA courses, such as *General Linguistics*. Additionally, they will use the *Guess the Language* game, developed within the learning block *Upskilling Your “Introduction to Language Variation” Course*, in the BA degree course, but probably not the entire learning block, as it is too advanced for the time available in their linguistics courses. The learning block *Automatic Speech Recognition and Forced Alignment* will also potentially be used in a module on technologies for interpreting within the MA degree course in *Interpreting*.

To this effect, during the first and second semesters of 2022-2023, under the supervision of **Maja Miličević Petrović**, MA students of *Specialized Translation* adapted and translated some of the learning blocks into Italian (*A Glimpse into Language Data Science, Collecting Language Data from Human Participants, Processing Texts and Corpora and Start Programming with Python in 10 Steps*), as well as the *Guess the Language* game from the *Upskilling Your “Introduction to Language Variation” Course* learning block.

## 2.3 CLARIN ERIC

CLARIN ERIC piloted parts of the learning block called *Introduction to Language Data: Standards and Repositories* at universities in the CLARIN network and during summer schools in corpus linguistics. More precisely, in the second semester of 2022-2023:

- A researcher/teacher from the Dutch Language Institute in the Netherlands (CLARIN centre) piloted the following thematic units in the 5 ECTS MA course *Computational Corpus Analysis* at Leiden University in the Netherlands:
  - 2.1 Introduction to research data repositories
  - 2.2 Metadata
  - 2.3 How repositories help make research data FAIR
  - 3.3 Finding tools in CLARIN to process digital text collections
- **Darja Fišer** piloted parts of the following thematic units in the course *Research Methods and Analysis Techniques in Digital Linguistics* within the PhD in *Linguistics* degree course in the Faculty of Arts at the University of Ljubljana in Slovenia:
  - 1.2. How are Language Resources Created, Managed and Shared?
  - 1.8. Create a Research Data Management Plan for Linguistic Research
- **Satu Salanski**, CLARIN ambassador and researcher/teacher, piloted parts of the materials during the *MEDAL Summer School in Corpus Linguistics*, which took place at the Institute of Estonian and General Linguistics at the University of Tartu in Estonia on 19-23 June 2023.
- **Martin Wynne**, representative of the CLARIN User Involvement Committee in the UK, adapted a tutorial handout produced in UPSKILLS during a pre-conference workshop on 2 July 2023 at the *Lancaster Corpus Linguistics* conference. The tutorial demonstrates how to search for linguistic corpora across several corpora collections hosted in different repositories using the CLARIN federated content search.

In the second semester of 2023-2024, teachers from Radboud University in the Netherlands plan to use *Analytical Thinking and Problem Solving* and *The Essence of Machine Learning for Linguists in Tech* to teach courses at the BA and MA levels. Furthermore, the teachers from the University of Pisa, Italy, plan to reuse *Introduction to Language Data:*

*Standards and Repository* in their BA-level Encoding Course, part of the Informatica Umanistica programme. CLARIN will continue to promote and disseminate the UPSKILLS learning blocks and related project deliverables after the end of the project to universities in Europe and beyond via the [Annual Conference](#), [User Involvement Committee](#) and [National Consortia](#). Furthermore, the learning content will be integrated into new European projects and adapted for different target audiences.

## 2.4 University of Geneva

During the first semester of 2023-2024, project members **Margherita Pallottino** and **Genoveva Puskas** will respectively use the following learning blocks:

- *Processing Texts and Corpora*, as part of the course *Introduction to Corpus Linguistics for students of English* (6 ECTS)
- *Upskilling Your “Introduction to Language Variation” Course* as part of the course *Introduction to Language Variation* (6 ECTS)

The courses will be open to BA and MA students of *English* or *Linguistics*.

## 2.5 University of Graz

In the second semester of 2022-2023, project members **Marko Simonović** and **Boban Arsenijević** piloted the thematic unit *Production Tasks in Second Language Acquisition* from the learning block *Collecting Language Data from Human Participants* in the 3 ECTS BA/MA course *Acquisition, Representation and Analysis of Data in Linguistics*. In the same semester, the project participants also piloted the thematic units *Word-Level Phenomena: Judgement Tasks* as well as *Word-Level Phenomena: Production Tasks* from the learning block *Collecting Language Data from Human Participants* in the 5 ECTS BA/MA course *The Linguistic Seminar in BCS and Slovenian*. They have also used various materials from the *Collecting Language Data from Human Participants* learning block in their supervision of BA theses.

Marko Simonović and Boban Arsenijević also planned to pilot the *First Steps into Scientific Research* and *Upskilling Your “Introduction to Language Variation” Course* learning blocks in the same semester. Unfortunately, the target course (a BA seminar) was cancelled due to a low number of registered students. They plan to do the piloting in the first semester of 2023-2024, as the same course will (most probably) take place then.

Our partners from the University of Graz plan to more systematically incorporate the UPSKILLS materials in their courses in the future. They are especially relevant and will have the biggest effect in the new cross-faculty 12 ECTS module called *Empirical Methods in Linguistics* that will be offered to the MA students of all degree courses involving languages and linguistics,

and which will consist of two courses: one with an introduction to various empirical methods in linguistics, and another in which students run their research projects. The decision to introduce the module was brought in one of the meetings of the Study Programmes forum, a series of events intended to lead to the modernisation of the degree courses in the Faculty of Humanities at the University of Graz, in which Marko Simonović gave a presentation of the UPSKILLS project and its results, with a focus on the *Guidelines and Best Practices for Research-Based Teaching* and on the learning content.

## 2.6 University of Malta

During the first and second semesters of 2022-2023, substantial parts of the following learning blocks were piloted in the BA in *Linguistics* degree course:

- *A Glimpse into Language Data Science*, as part of the LLT2810 *Introduction to Research Methods in Linguistics* (6 ECTS) study unit by **Paul Marty**
- *Processing Texts and Corpora*, as part of the LIN3095 *Corpus Linguistics* (4 ECTS) study unit by **Stavros Assimakopoulos**
- *Start Programming with Python in 10 Steps*, as part of the ICS1251 *Programming In Python: A Practical Introduction* (6 ECTS) study unit by **Marc Tanti**
- *The Essence of Machine Learning for Linguists in Tech*, as part of the LIN2031 *Machine Learning for Natural Language Processing* (6 ECTS) study unit by **Marc Tanti**

Following the success of this piloting exercise, project participants plan to permanently incorporate the relevant materials into their teaching in these study units. In addition, materials and exercises from other UPSKILLS learning blocks will be used in the following study units:

- LIN1024 - *Hands-on Language* (4 ECTS), to include materials from *First Steps into Scientific Research* and *Analytical Thinking and Problem Solving*
- LLT1160 - *Sociolinguistics: An Introduction to Language Variation and Change* (4 ECTS), to include materials from *Collecting Language Data from Human Participants*
- LIN3010 - *Language Typology and Universals* (4 ECTS), to include materials from the *Upskilling Your “Introduction to Language Variation” Course*
- LIN5011 - *Advanced Research Methods in Linguistics 1* (6 ECTS), to include materials from *Introduction to Language Data: Standards and Repositories*

## 2.7 University of Rijeka

Three lecturers, two of whom are project members, in the Faculty of Humanities and Social Sciences are interested in using the UPSKILLS learning content. Project member **Tihana Kraš**, teaching first language acquisition, second language acquisition, bilingualism and

multilingualism, plans to use the following three learning blocks to teach 3 ECTS courses *Second Language Acquisition* and the *Acquisition of English as a Second Language* within the teaching track of the MA degree course in *English Language and Literature* in the first and second semester of 2023-2024 and in the years to follow:

- *A Glimpse into Language Data Science*
- *Collecting Language Data from Human Participants*
- *First Steps into Scientific Research*

In the second semester of 2022-2023, four thematic units of the *Collecting Language Data from Human Participants* learning block were successfully piloted within the 3 ECTS MA course *Acquisition of English as a Second Language: Comprehension Tasks in Second Language Acquisition, Production Tasks in Second Language Acquisition, Acceptability Judgment Tasks in Second Language Acquisition* and *Ethics in Linguistic Research with Human Participants*.

Project member **Martina Podboj**, teaching English phonetics and phonology, discourse analysis and language testing and assessment, plans to use parts of the *Collecting Language Data from Human Participants* learning block to teach her courses in the future within the BA and MA degree courses in *English Language and Literature*. A **lecturer** teaching digital humanities, computational linguistics, cognitive linguistics, data science and Python programming plans to use the following learning blocks to teach his courses at the BA and MA level in the first and second semester of 2023-2024:

- *A Glimpse into Language Data Science*
- *Analytical Thinking and Problem Solving*
- *Automatic Speech Recognition and Forced Alignment*
- *Collecting Language Data from Human Participants*
- *First Steps into Scientific Research*
- *Introduction to Language Data: Standards and Repositories*
- *Processing Texts and Corpora*
- *Project Management*
- *The Essence of Machine Learning for Linguists in Tech*

In the Faculty of Maritime Studies, a **lecturer** teaching English for specific purposes and translation studies plans to use the following learning blocks to teach the course *Language Technologies for Developing Written Communication* within the PhD degree course in *Maritime Studies* and to teach lifelong learning courses in the second semester of 2023-2024:

- *A Glimpse into Language Data Science*
- *Introduction to Language Data: Standards and Repositories*
- *Start Programming with Python*
- *Processing Texts and Corpora*

- *The Essence of Machine Learning for Linguists in Tech*

### **3. Dissemination outside the Partner Institutions**

Via two surveys we sent out in November 2022 and July 2023, lecturers from **six** universities outside the UPSKILLS consortium expressed interest in integrating the UPSKILLS learning content into their university curricula and using the guidelines developed in the project in the next two academic years (2023-2024 and 2024-2025), as outlined below.

#### **3.1 University of Vienna, Austria**

A lecturer teaching audiovisual translation and accessibility is interested in using the following learning blocks and guidelines in his/her teaching at the BA level in the first semester of 2023-2024:

- *A Glimpse into Language Data Science*
- *First Steps into Scientific Research*
- *Processing Texts and Corpora*
- *Guidelines and Best Practices for Research-Based Teaching*

Another lecturer teaching translation technologies, machine translation and research skills is interested in using all the UPSKILLS learning blocks, except *Start Programming with Python in 10 Steps*, in the first semester of 2023-2024.

#### **3.2 KU Leuven University, Belgium**

A lecturer teaching applied linguistics is interested in using the following learning blocks in her teaching at the BA and MA level in an unspecified period, which depends on the schedule of the courses and the curriculum flexibility:

- *First Steps into Scientific Research*
- *Text Processing*
- *Introduction to Language Data: Standards and Repositories*
- *Start Programming with Python in 10 Steps*
- *The Essence of Machine Learning for Linguists in Tech*

#### **3.3 University of Osijek, Croatia**

In the Faculty of Humanities and Social Sciences, a lecturer teaching psycholinguistics and semantics plans to use the following learning blocks to teach courses *Writing Academic Papers in Linguistics*, *Introduction to Sociolinguistics* and *Psycholinguistics* within the BA and MA

degree courses in *English Language and Literature* in the first and second semester of 2023-2024 and 2024-2025:

- *Analytical Thinking and Problem Solving*
- *Collecting Language Data from Human Participants*
- *First Steps into Scientific Research*
- *Upskilling Your “Introduction to Language Variation” Course*

### **3.4 University of Zagreb, Croatia**

In the Faculty for Teacher Education, a lecturer teaching linguistics, Croatian language and the early learning of Croatian as a mother tongue plans to use the following learning blocks to teach courses about (a) the Croatian language, (b) public speaking and (c) language games at the BA, MA and PhD level in the second semester of 2023-2024 and the first semester of 2024-2025:

- *A Glimpse into Language Data Science*
- *Analytical Thinking and Problem Solving*
- *Collecting Language Data from Human Participants*
- *First Steps into Scientific Research*
- *Introduction to Language Data: Standards and Repositories*
- *Processing Texts and Corpora*
- *Upskilling Your “Introduction to Language Variation” Course*

In the same faculty, a lecturer teaching Croatian and second language acquisition plans to use the same learning blocks to teach the courses *L1 and L2 Acquisition: Applied Linguistic and Neurolinguistic Perspectives*, and *Educational Linguistics* within the PhD degree course *Linguistic, Literary and Cultural Context of Early, Preschool and Primary Education* in her home institution in the first and second semester of 2023-2024 and perhaps in 2024-2025.

### **3.5 Vytautas Magnus University, Lithuania**

A lecturer teaching translation studies, corpus linguistics, and academic writing is interested in using the following courses and guidelines at the MA level in the second semester 2023-2024:

- *Analytical Thinking and Problem Solving*
- *First Steps into Scientific Research*
- *Processing Texts and Corpora*
- *Introduction to Language Data: Standards and Repositories*
- *Guidelines for Using Research Infrastructures for Language Resources and Technologies*

### 3.6 University of Ljubljana, Slovenia

A lecturer teaching Japanese linguistics and translation is interested in using the following courses and guidelines at the MA level in the first semester of 2023-2024:

- *Introduction to Language Data: Standards and Repositories*
- *Text Processing*
- *Guidelines for Using Research Infrastructures for Language Resources and Technologies*
- *Guidelines for Creating Learning Content*

☞ All the lecturers interested in using the UPSKILLS learning content will be invited to create an account on Moodle to be able to view and browse the materials. Access to Moodle is facilitated via the [project website](#), which contains the summary of each learning block. On Moodle, all learning blocks share two tiles in common that are relevant for those using our materials:

1. *About This Learning Block*: This tile provides instructions about how to use the learning content.
2. *You Like What You See*: This is the last tile in each block providing the citation format and licence information, as well as a Moodle .mbz file with the content that can be downloaded and uploaded to another Moodle system. Additionally, the zip files will be archived in a public repository for long-term preservation and sharing.

## 4. Dissemination via Open Registries and Repositories

As part of UPSKILLS Task 3.4, one of the subtasks is to present a list of relevant open registries and repositories potentially suitable for disseminating the UPSKILLS learning content. For clarification purposes, a **registry** usually represents a central location managed by an organisation that registers and maintains the metadata describing the content of learning resources hosted in separate locations, e.g. [SSH Open Marketplace](#) or [CLARIN Learning Hub](#). A registry is useful for making training and learning resources findable. On the other hand, a **repository** can host both the metadata and the content of the learning resources. Moodle, for example, is also a repository. However, its access can be restricted by an organisation and its sustainability is usually not guaranteed. Therefore, in this section, we highlight a few public repositories that can be used for the dissemination, archiving and preservation of the UPSKILLS learning content in the long term. The identified registries and repositories are summarised here,

along with submission guidelines, to help learning block creators get acquainted with the overall requirements and process. Additionally, this section clarifies which repositories are suitable in which scenarios to include the UPSKILLS learning content.

## 4.1 The Digital Humanities Course Registry

The [Digital Humanities Course Registry \(DHCR\)](#) is a free online registry jointly developed and maintained by the CLARIN and DARIAH research infrastructures. Lecturers in DH can use the platform to promote their programme, courses or other curricula DH-related activities. At the same time, students can search for upskilling opportunities either in their own country or abroad.

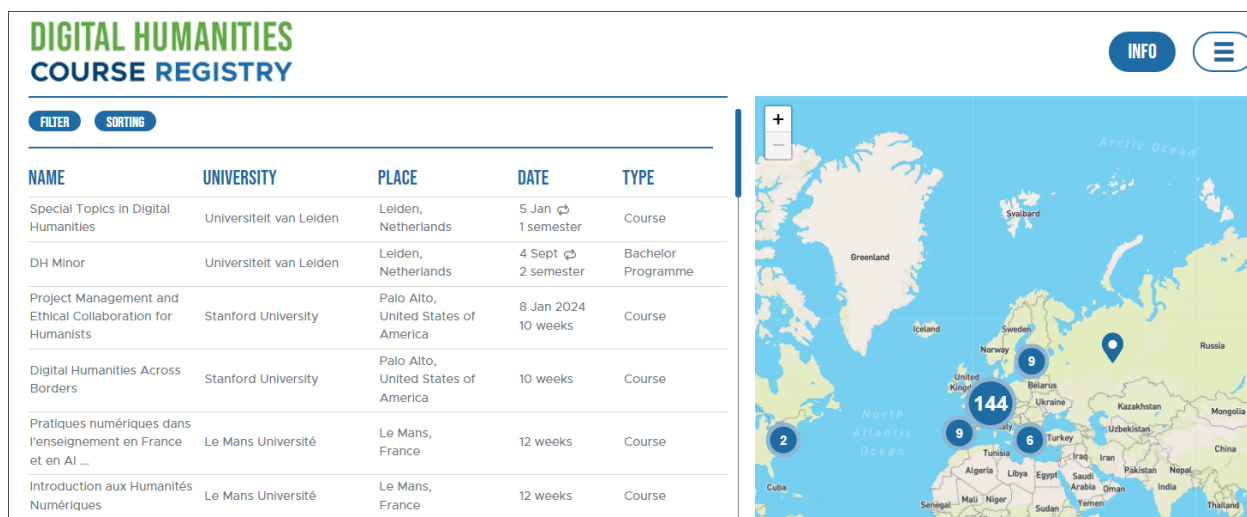


Figure 1. Browsing interface of the Digital Humanities Course Registry

The following types of courses can be included in the registry: on-site and online courses, modules, whole programmes (BA, MA, PhD), summer schools and workshops for professional development. Both one-off and recurring educational courses can be added to the registry. The main requirement is that the courses are part of university curricula, require enrollment, and have a clear start and end date. Once the course metadata is submitted to the registry, the [national moderator](#) reviews and approves the entry. All course contributors are automatically reminded to check and update their courses' metadata at least once a year.

While the DH Course Registry can help spread awareness about the UPSKILLS learning blocks within the DH community, it is not appropriate to feature the materials in the registry

because they are designed for self-paced, asynchronous learning. Therefore, the UPSKILLS materials are better suited for inclusion in the [DARIAH CAMPUS](#), see section 4.3. However, if the consortium partners plan to reuse the UPSKILLS learning content as part of specific programmes at their university curricula, they could contribute the metadata of those programmes or courses to the DH course registry, mentioning the UPSKILLS project in the course description. In such a case, each consortium partner can add the metadata of their programme (s) to the DH Course Registry in the following way:

1. Fill in the DHCR [User Registration form](#) to create an account.
2. Once the administrator approves the account, access the dashboard, go to **Administrate courses** and click on **Add course**. Add the following metadata of your UPSKILLS learning block:
  - a. Enter the name of the learning content block.
  - b. Indicate whether the course is offered **online**.
  - c. Select education type.
  - d. Select language: **English**
  - e. Indicate the number of ECTS.
  - f. Add the hyperlink to your course or programme.
  - g. Add the name and the email address of the main contact person.
  - h. Add a short description of your course, including learning outcomes and specifying which UPSKILLS units or blocks have been integrated into the course or programme.
  - i. Entry requirements: specify if the learner needs specific background knowledge to enrol in your programme.
  - j. Add the start date: the publication date and indicate that it is a recurring course
  - k. Duration: add **1**
  - l. Duration type: **semester**
  - m. Add your institution
  - n. Add your department
  - o. Disciplines (you can select more than one)
  - p. Select the research methods and objects taught in the programme. Note that the research methods are based on [TaDiRAH](#) taxonomy, which targets research methods and tools in the digital humanities.

When done, tick the box **Show course in the registry**. The metadata will be reviewed by the national DHCR moderator and published in the database. If a country does not have a national moderator, one of the DHCR administrators will approve the entry.

## 4.2 The SSH Open Marketplace

The [SSH Open Marketplace](#) was developed within the framework of the [Social Sciences & Humanities Open Cloud \(SSOC\) project](#) to provide support for researchers at every step of the research data lifecycle through user-friendly discovery services of existing tools, datasets, training materials, workflows and publications. Please note that the marketplace does not host resources, just links to them. Although the project ended in 2022, the consortium partners committed to maintaining the platform through a memorandum of understanding. All the entries created in the Marketplace are first curated by content editors and then published. A citation format is provided as well. Here is an example of a tutorial published on the CLARIN website and made discoverable via the marketplace: [Jupyter notebooks for Europeana newspaper text resource processing with CLARIN NLP tools | Social Sciences & Humanities Open Marketplace \(sshopencloud.eu\)](#). The European Open Science Cloud may harvest the metadata of the Marketplace entries in the near future.

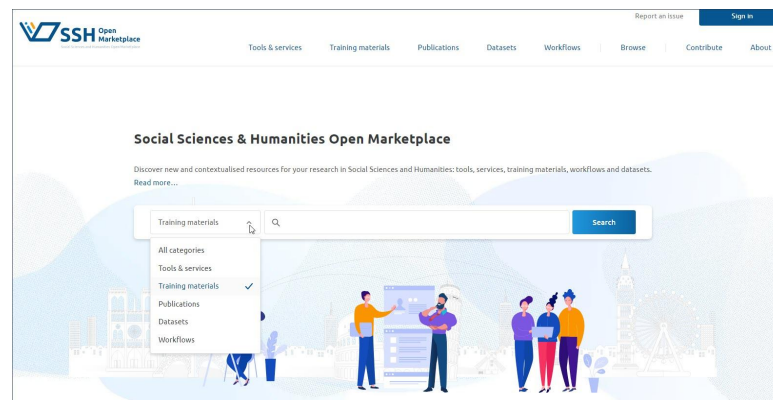


Figure 2. The Search Interface of the SSH Open Marketplace

The criteria for proposing new training materials are:

- The material needs to be **relevant** to academics, scholars and students from the SSH field and in line with the digital methodologies used within the SSH landscape.
- The material should come from a primary source and be current and **open**.
- The **metadata describing the material should match the metadata model** used in the SSH Marketplace.
- The material should comply with GDPR regulations.

Evaluating the above **inclusion criteria**, the UPSKILLS courses on Moodle seem to be a suitable training resource for the marketplace as they aim to enhance students' research skills

from language and linguistics studies. However, the Moodle platform is not completely open because it requires account registration and enrollment into a specific learning block. Supposing that this is not a barrier, each consortium partner can add the metadata of their UPSKILLS learning content blocks in the following way:

1. Sign in with EOSC<sup>1</sup> credentials using existing accounts such as Google, DARIAH, eduTeams, or academic accounts.

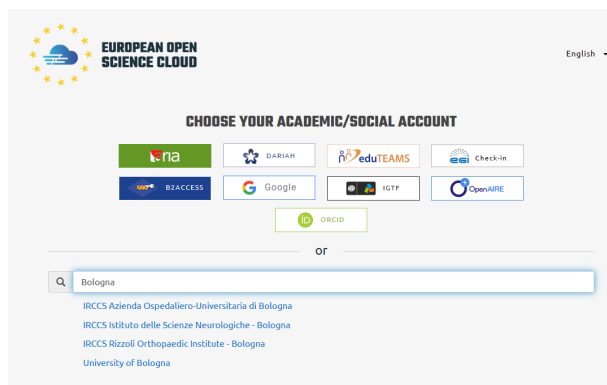


Figure 3. How to Access the SSH Open Marketplace

2. Click on **Create Training Material** and add the mandatory and recommended metadata, e.g.
  - a. Name (of the unit block/material)
  - b. Version
  - c. Description (concise)
  - d. Accessible at (URL, landing page)
  - e. Date created
  - f. Date last updated
  - g. Actors (e.g. Author, Contributor, Reviewer)
  - h. Property > Activity (e.g. data analysis, programming)
  - i. Property > Keywords (e.g. data analysis, data management, data discovery, data processing)
  - j. Property > Discipline (see the [Austrian Fields of Science and Technology Classification 2012](#))
  - k. Property > Language
  - l. Property > Intended audience (select *Trainer*; see [SSHOC audience concept scheme](#))

---

<sup>1</sup> European Open Science Cloud

- m. Property > Licence (select CC-BY4.0 International)
- n. Property > Terms of use URL
- o. Property > Year of publication
- p. Property > Resource category (select *Education*)

Once submitted, the entry will be curated by the platform editors and approved. More information can be found on the SSH Marketplace website - [Guidance for metadata fields](#).

Besides the UPSKILLS courses, the consortium partners can also add the metadata of any datasets, publications, and/or tools used in the course, which are publicly available.

### 4.3 DARIAH Campus

[DARIAH-Campus](#) is a discovery and hosting platform for open educational resources produced for **asynchronous** learning in Arts and Humanities. It has been developed and maintained by DARIAH-EU, a European research infrastructure for arts and humanities scholars. Other infrastructures, such as [EHRI](#), have also started using the platform to include their learning resources.

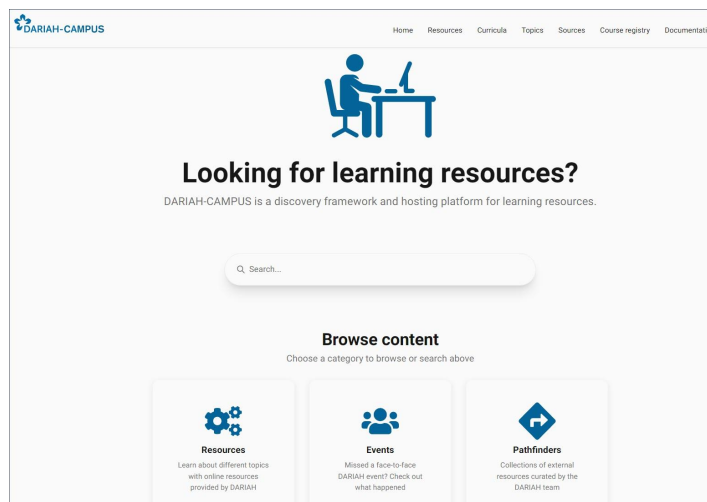


Figure 4. The Search Interface of DARIAH-CAMPUS

The platform can be used for several purposes:

1. **To link to external learning resources** hosted on other sites, created by individuals or groups based at institutions that have a [connection](#) with DARIAH-EU, the national nodes or working groups. After contacting the platform owners, we understood that the connection to DARIAH-EU is not a mandatory requirement to be able to contribute to the platform. Example of an external resource: [Introduction to Programming for NLP with Pyhton](#), hosted by the University of Bergen.

2. **To link to hosted resources** developed by trainers using the Content Management System and GitHub offered by DARIAH-Campus. Example of a hosted resource: [Data Protection in Research Practice](#).
3. **To create pathfinders, i.e.** a hybrid resource developed by DARIAH-CAMPUS, which guides the learner through a specific topic, e.g. [DARIAH Pathfinder to Data Management Best Practices in the Humanities](#) or the [ELEXIS Pathfinder to Computational Lexicography for Developers and Computational Linguists](#).
4. **To upload learning resources** produced during training webinars, lectures or workshops on specific topics. E.g. [CLS-INFRA Training School on Data and Annotation](#).

All the learning resources shared or hosted via the platform can be **cited** and **reused** per the [DARIAH-Campus Reuse Charter](#), and are also ingested in the SSH Open Marketplace.

As the UPSKILLS learning content is hosted on Moodle, it would qualify as an **external resource**. However, the content can be linked to the DARIAH-Campus only if it is **openly available for use and reuse**, see [example](#). In our case, the content is available for use and reuse, but registration on Moodle is required. It remains to be investigated whether this impedes submitting a resource. Additionally, DARIAH-Campus provides a sustainable alternative for any consortium partner who intends to update or manage their learning content outside the Moodle platform after the end of the UPSKILLS project. For instance, the CLARIN research infrastructure, a partner in UPSKILLS, does not have a Moodle system. Therefore, CLARIN seeks a platform to maintain its learning blocks (*Introduction to Language Data and Automatic Speech Recognition*) and customise them for new target audiences. If CLARIN decides to host their resources on DARIAH-Campus, some extra effort would be required to prepare and transfer the materials from Moodle to GitHub and make them editable in Markdown. The multiple-choice quizzes may need to be imported or developed from scratch, depending on the level of interoperability between Moodle and the DARIAH CMS. Please refer to the [Documentation](#) for more detailed instructions on using the [Content Management System](#) and [GitHub](#).

#### 4.4 OER Commons

[OER Commons](#) is a public digital library of **open educational resources**, which UNESCO defines as “learning, teaching and research materials in any format and medium that reside in the public domain or are under copyright that have been released under an open license, that permits no-cost access, re-use, re-purpose, adaptation and redistribution by others”. The following types of collections can be shared via the OER Commons:

- Full university courses
- Interactive mini-lessons and simulations

- Adaptations of existing open work
- Open textbooks
- Lesson plans, worksheets, and activities

Resources from all subject fields and all education levels can be added. A learning resource can be added to OER if it meets the following conditions:

- It supports learning in undergraduate, graduate, vocational, or professional learning.
- It provides a clear context or use for a specific educational audience.
- It contains articulated learning goals and can be aligned to educational standards.
- It is released within the public domain under a Creative Commons licence.
- It is openly accessible without creating an account or logging in.
- It does not require the use of commercial software or services.
- It is appropriately granular to facilitate ease of use.
- The software required to use the resource is supported in modern operating systems and browsers.

OER Commons does not only link to open educational resources, but it also provides tools (e.g. Open Author) to build and publish new ones.

According to the UNESCO definition, the UPSKILLS learning content can be described as OER and disseminated via OER platforms. The UPSKILLS materials meet most of the criteria listed above, except that Moodle access to browse the resources requires users to create an account. Assuming this does not pose a barrier, the UPSKILLS learning blocks could be included in OER Commons by following the steps below:

1. Create an OER account by clicking on **Add OER**.
2. Select the option to **submit a link from the Web**.
3. Add the URL to the [summary of a learning block](#) on the project website.
4. Name and describe the resource.
5. Attribute the resource by entering all the authors.
6. Select the licence: [CC BY 4.0](#).
7. Add information to make the resource discoverable by others: subject areas, education level, material type, languages, etc.
8. Add optional information, e.g. educational use, primary user, accessibility, and tags.
9. Preview and submit resources.

For more information and instructions, please see the [Submission Guidelines](#). By including the UPSKILLS materials in OER Commons,

## 4.5 Zenodo

Zenodo (<https://zenodo.org>) is an open repository for researchers, scientific communities and academic institutions from all disciplines to share and preserve their research outputs and educational materials. Different data types in any file format and size can be uploaded, shared and accessed for free. A unique identifier (DOI) is automatically assigned to each upload to make it citable and harvestable by third parties. Furthermore, learning content developed on GitHub can obtain a DOI by sharing their content with the Zenodo repository, see [this guide](#) on the Sage blog.

Zenodo has already been selected to disseminate the general project deliverables, please see the [UPSKILLS project](#) created in Communities. Besides general project outputs, the repository could also be used to deposit and disseminate the UPSKILLS learning content in the wider SSH communities. Uploading learning materials can be achieved in a few easy steps:

1. Go to Communities and search for the [UPSKILLS project community](#), which has been specifically created to disseminate the project deliverables.
1. Access Zenodo with your ORCID or GitHub ID or create a new account.
2. Click on **Upload** to start a new entry.
3. Drag & drop the files (max. 50GB per dataset). It would be useful to include a ReadMe file along with the Moodle .mbz file, which contains information about how others can unzip the file and reuse the learning content in their own Moodle system. Of course, the authors can also opt to upload and share only specific parts of their learning content, such as a PowerPoint presentation or a handout in PDF format.
4. Select the upload type: Other<sup>2</sup>
5. **Reserve a DOI** for your upload.
6. Add the required basic information, e.g.
  - The description should include information about the UPSKILLS project, respective learning block, nr of ECTS, target audience, the main learning outcomes, and a few instructions on how others can use the learning content in their university curricula.
  - Add keywords to describe the learning content.
7. Select the licence > Open Access > CC BY 4.0.
8. Add the UPSKILLS Erasmus funding grant.
9. Add optional information, e.g. contributors.
10. Publish the entry and share it with others.

---

<sup>2</sup> This is the best option to upload all your course files in one go. To share only one lesson, the *Lesson* type needs to be selected. In this case, it is recommended to include the learning outcomes for that specific lesson in the Zenodo description and instructions to make it reusable by other teachers.

## 4.6 CLARIN.SI Repository

The CLARIN.SI repository is a research data repository hosted by the CLARIN Slovenian infrastructure. Unlike Zenodo, which caters to all disciplines and accepts all types of research outputs, the CLARIN repositories specifically target those disciplines that involve language and linguistic research. Although the repositories are mainly used for publishing and disseminating language resources (e.g. corpora) and technologies, they can also be used to publish and archive training materials produced for a project in which CLARIN participates. Each entry in the CLARIN repositories is assigned a persistent identifier (handle), which can be used for citation purposes and to make all resources automatically discoverable via the [Virtual Language Observatory](#), a free metadata catalogue. Uploading learning materials to the repository can be achieved as follows:

1. Log in using your CLARIN account or institutional credentials.
2. Read the [depositing guidelines](#).
3. Then go to your **Account** and access the [Submissions](#) area.
4. Click on **Create a submission**. When prompted to select the resource type, please select **toolService**.
5. Fill in the required metadata, such as title, project URL, demo URL, authors, publisher (e.g., institution), contact person, funding information, description of the learning content and how it can be reused, language, and keywords best suited to describe your content.
6. Drag and drop the files (e.g. 01\_Unit 1\_Introduction.mbz)
7. Read and accept the [Distribution License Agreement](#) and select the licence (CC-BY 4.0)
8. **Save & Share** the submission with the repository manager.

After the repository manager has approved the submission, the learning content will be findable via the CLARIN.SI repository and the Virtual Language Observatory. As CLARIN is one of the consortium partners in UPSKILLS, this repository seems the most sustainable solution for archiving the UPSKILLS learning content long-term. For more information about research data repositories, please refer to the [Integration of Research Infrastructure into Teaching: Recommendations and Best Practices](#) ( van der Lek, I., Fiser, D. et al., 2023).

To summarise, section 4 includes an overview and brief analysis of open registries and repositories suitable for disseminating the UPSKILLS learning materials at the end of the project. OER and Zenodo cater to all disciplines, while DH Course Registry, DARIAH-Campus, SSH Open Marketplace and CLARIN.SI serve specific communities in the Social Sciences, Arts and (Digital) Humanities. Although the DH-Course Registry may be used to raise awareness of the UPSKILLS learning content in the DH community, the platform accepts only metadata about courses and training activities designed for synchronous learning. The only solution to include

the UPSKILLS materials in the DH Course Registry is to integrate them into an official university course or programme. Further, while the SSH Open Marketplace and OER Commons are suited for dissemination purposes, Zenodo and CLARIN.SI repositories are reliable solutions for archiving the UPSKILLS learning content for long-term preservation and reuse. Each repository assigns a unique identifier to the entries, e.g. DOI and handle. Finally, for those interested in maintaining their learning block outside Moodle after the end of the project, DARIAH Campus is a viable solution. However, this would require additional investment in terms of time and effort.

The criteria for publishing learning and teaching materials vary depending on the specific platform; however, they all share common criteria for sharing the materials:

1. **Open licence and Terms of Use:** The materials need to be published online, preferably fully open (no user account needed), under a non-restrictive Creative Common Licence or GNU General Public Licence, and specifying the Terms of Use. If the UPSKILLS learning content is based on other public learning resources, the resources must be properly referenced according to the licence and terms of use. If the reused materials are not published with a licence, the authors should be informed that they have been adapted for UPSKILLS and will be released under a CC BY 4.0 licence.
2. **Educational value:** The materials must have a clear educational purpose, include learning outcomes and target a specific audience.
3. **High quality, modular content:** The materials should be well-organised, as granular as possible and easy to understand. For example, authors should consider preparing and uploading zip files of individual units within a specific learning block, which would increase its reusability potential.
4. **Technical compatibility:** The materials should be shared in common file formats, e.g. text files, MS formats, markdown. The Moodle zip packages are generally accepted for archiving in Zenodo and CLARIN repositories. Still, clear instructions should be provided on reusing the zip files in other Moodle systems. The same applies to learning content and games produced in H5P format.

## 5. Dissemination plan and materials

Once most of the learning blocks have been completed, we will open up Moodle and create an overview of the learning blocks on the project website with access to Moodle. We will also include an invitation to browse our courses on Moodle in the UPSKILLS newsletter. Additionally, we will draft an email template presenting the project and the learning content on Moodle and invite partners to send it to the relevant contacts (individuals, institutions, associations, etc.) in their national context and further. This email template contains an invitation to the relevant contacts to use the UPSKILLS learning content in their teaching and integrate it into their curricula and the link to a questionnaire in the Google Forms format in which they will register their intention to do so.<sup>3</sup> Finally, to disseminate the learning content as much as possible, we will develop additional promotional materials that can be used to this effect. These include a **press release**<sup>4</sup> marking the conclusion of the project, a **slide deck**<sup>5</sup> for a 1.5 to 2-hour presentation that covers the whole project (which project partners can use and adapt as needed), a **digital brochure**<sup>6</sup> for the overall project, **flyers**<sup>7</sup> for each one of the learning content blocks and **four promotional videos**. The videos will include:

- An introduction to the overall project
- A walkthrough demo of the website, guides and learning content on Moodle
- Interviews with the students attending our Summer School in Petnica about their experience with working on small-scale research projects
- Interviews with the lecturers teaching at the Summer School

The press release and the slide deck will be circulated among partners. The brochure and the flyers will be uploaded on [the project website](#), and the promotional videos will be published on our [UPSKILLS YouTube channel](#), and linked to the [CLARIN channel](#) for wider dissemination.

## Conclusions

This document outlined the plan for disseminating the UPSKILLS learning content within and beyond the partner institutions. According to the data collected via two surveys, we have established that the learning content will be integrated into the curricula of six partner institutions

---

<sup>3</sup> August 2023 update: This e-mail template is now available in [Annex A](#).

<sup>4</sup> August 2023 update: This press release is now available in [Annex B](#).

<sup>5</sup> August 2023 update: This slide deck is now available in [Annex C](#).

<sup>6</sup> August 2023 update: This brochure is now available in [Annex D](#).

<sup>7</sup> August 2023 update: All flyers are now available in [Annex E](#).

(funded by the EU and Movetia) and the curricula of six universities outside the UPSKILLS consortium. We have also identified registries and repositories for open educational resources that would be potentially suitable for including the UPSKILLS learning content and outlined their requirements. Among these, we have recommended adding the metadata of the UPSKILLS learning content blocks to the SSH Open Marketplace<sup>8</sup> to make the materials discoverable in Social Sciences and Humanities communities. For archiving the UPSKILLS learning content for the long term, we recommend using either Zenodo (UPSKILLS Community) or the CLARIN.SI<sup>9</sup> repository. This way, the materials will be discoverable and reusable even if the Moodle platform is no longer maintained after the end of the project. Additionally, we have developed a plan for opening up Moodle and inviting individuals, institutions, associations and others to browse its content and use it. Finally, we have devised a plan for developing and disseminating additional promotional materials (a press release, an email template, a registration form for using the UPSKILLS learning blocks, a PPT presentation, a press release, digital brochures, flyers and promotional videos).

---

<sup>8</sup> August 2023 update: The metadata of the learning content is now available in SSH Open Marketplace: <https://marketplace.sshopencloud.eu/training-material/yweKFs>.

<sup>9</sup> August 2023 update: The learning content blocks are now available for download from CLARIN.SI: <http://hdl.handle.net/11356/1865>.

## References

- CLARIN.SI Repository. *How to Deposit*. Retrieved 15 August 2023. <https://www.clarin.si/repository/xmlui/page/deposit>.
- DARIAH-Campus. Documentation. Proposing a Contribution. Retrieved 15 August 2023. <https://campus.dariah.eu/docs/proposing-a-contribution>
- Digital Humanities Course Registries. How to Use DHCR. Retrieved 15 August 2023. <https://dhcr.clarin-dariah.eu/info#purpose>
- Methodspace. How to get a DOI for your teaching materials with Zenodo. Retrieved 15 August 2023. <https://www.methodspace.com/blog/how-to-get-a-doi-for-your-teaching-materials-with-zenodo>
- Open Educational Resources. Submission Guidelines. Retrieved 15 August 2023. <https://help.oercommons.org/support/solutions/articles/42000046850-submission-guidelines>
- SSH Open Marketplace. Create an individual item. Retrieved 15 August 2023. <https://marketplace.sshopencloud.eu/contribute/create-an-individual-item>
- van der Lek, Iulianna, Fišer, Darja, et al. (2023). Integrating Research Infrastructures into Teaching: Recommendations and Best Practices. The Case of CLARIN. UPSKILLS Intellectual Output 2.2. Project Deliverable. <https://doi.org/10.5281/zenodo.8114407>

## Annex A: UPSKILLS dissemination e-mail template

### ENGLISH VERSION

Dear colleagues,

[UPgrading the SKills of Linguistics and Language Students \(UPSKILLS\)](#) is a strategic partnership for higher education between institutions (primarily universities) from Austria, Croatia, Italy, Malta (the coordinator), the Netherlands, Serbia and Switzerland funded by European Union's Erasmus+ and Swiss Movetia. The project seeks to identify and tackle the gaps in skills of linguistics and language students in order to enhance their employability and make them better prepared for careers in research and industry.

To achieve this goal, the project partners are developing open-access teaching and learning materials to be embedded in existing and future university curricula, best practices for research-based teaching and educational games. The open-access teaching and learning materials, organised in [eleven learning blocks](#), cover the areas of analytical thinking and problem solving, automatic speech recognition, language data collection from human participants, language data science, language typology and language variation, machine learning, project management, programming with Python, research methodology and text processing, and are intended to improve students' research, data manipulation, technical, organisational, transversal and disciplinary skills.

The learning blocks are now available for public view on the [Moodle platform](#); to access it, you need to create an account first. Please note that the learning blocks are still under review and will be finalised by the end of August, when the project ends. Teachers and researchers who wish to use some or parts of learning blocks in the next two academic years are kindly invited to let us know via this [registration form](#). If they would like to provide any feedback on the content after they have finished using it, they can email the learning blocks' creators, whose addresses have been provided on the Moodle platform.

If you have any questions, do not hesitate to contact us/me at XXXXXXXXXXXX.

Thank you very much for your interest in the UPSKILLS project!

Kind regards,

XY on behalf of the UPSKILLS consortium

## **Annex B: UPSKILLS conclusion press release**

### **PRESS RELEASE**

For Immediate Release

[Place], 30 August 2023

### **UPSKILLS Project aimed at enhancing the employability of graduates from linguistics and language-related courses has successfully been completed**

UPSKILLS (UPgrading the SKILLS of Linguistics and Language Students), an Erasmus+ partnership project, has successfully drawn to an end on the 31st of August 2023. This 3-year project, launched in September 2020, sought to identify and tackle the gaps and mismatches in skills for linguistics and language students. To reach this aim, consortium members from 8 partner institutions worked towards a new curriculum component which targets the development of technical and transferable skills needed for industry and academic research among undergraduate students of language-related subjects. This component, which has now been finalised, involves online and blended learning and relies on the use of innovative pedagogies, primarily online educational games, some of which have been developed specifically for this project.

The development of this new study component revolved around the production of four intellectual outputs:

1. In-depth needs analysis that gives rise to a detailed professional profile description for the relevant graduates, with the associated learning objectives (knowledge, skills and competences), typical tasks and responsibilities.
2. Identification of best practices and guidelines for integrating research into teaching.
3. Adaptation of existing - and, where necessary, creation of new - learning content, aimed at developing the technical and transferable skills of students of language-related disciplines.
4. Integration of educational games into the learning content in the interest of enhancing student engagement.

All of the deliverables published in relation to the above intellectual outputs are available on the [UPSKILLS website](#), together with the learning content which can additionally be browsed through our dedicated [Moodle platform](#).

Since the project's ultimate goal was to ensure that students in language-related disciplines have better job prospects and transition smoothly to the world of work, therefore creating a better workforce for industry and research, the project strove to engage the relevant students, academics and industrial stakeholders alike. To this end, a total of 6 multiplier events and 3 training events were organised as well. The project culminated in a [five-day summer school](#) – the third and final training event – held at the Petnica Science Centre, Serbia in July 2023. This summer school brought together a total of 32 students and 17 trainers from the University of Belgrade, University of Bologna, University of Geneva, University of Graz, University of Malta, University of Rijeka and CLARIN ERIC and engaged the students through hands-on and research-oriented lectures and workshops.

*[Contribution of your institution to the project...]*

For more information visit the UPSKILLS website: <https://upskillsproject.eu/>





### 3.1 A glimpse into language data science

This learning block introduces you to the field of language data science, which is an interdisciplinary area of research that combines linguistics, computer science, and statistics to analyze large amounts of language data.

- Objectives:**
1. Understand the scope and goals of language data science.
  2. Identify the key challenges and opportunities in this field.
  3. Explore the role of technology in language data science.
  4. Discuss the ethical implications of working with large-scale language data.

### 3.2 Analytical thinking and problem solving

This learning block focuses on developing analytical thinking and problem-solving skills. It includes exercises and case studies that challenge you to apply your knowledge to real-world scenarios.

- Objectives:**
1. Apply analytical thinking to solve complex problems.
  2. Develop effective problem-solving strategies.
  3. Work in teams to tackle challenging tasks.
  4. Present and defend your solutions.

### 3.3 Automatic speech recognition and forced alignment

This learning block explores the intersection of linguistics and computer science, specifically focusing on automatic speech recognition (ASR) and forced alignment. You will learn how these technologies are used in various applications.

- Objectives:**
1. Understand the principles of ASR and forced alignment.
  2. Analyze the challenges and solutions in this field.
  3. Explore the role of machine learning in ASR.
  4. Discuss the impact of ASR on language learning and research.

### 3.4 Collecting language data from human participants

This learning block discusses the methods and ethics of collecting language data from human participants. It covers the design of experiments, data collection techniques, and the importance of informed consent.

- Objectives:**
1. Design and conduct experiments to collect language data.
  2. Ensure ethical standards are followed throughout the process.
  3. Analyze and interpret the collected data.
  4. Discuss the challenges and limitations of human data collection.

### 3.5 First steps into scientific research

This learning block provides an overview of the scientific research process, from identifying a research question to presenting findings. It includes practical advice on how to approach a research project.

- Objectives:**
1. Identify a research question and formulate a hypothesis.
  2. Design a research plan and collect data.
  3. Analyze the data and draw conclusions.
  4. Present your findings and discuss their implications.

### 3.6 Introduction to language data: Standards & repositories

This learning block introduces you to the standards and repositories used in language data science. It covers the importance of data quality and the role of these resources in the field.

- Objectives:**
1. Understand the importance of standards and repositories.
  2. Identify key standards and repositories in the field.
  3. Discuss the challenges and solutions in this area.
  4. Explore the role of these resources in language data science.

### 3.7 Processing texts and corpora

This learning block focuses on the techniques for processing large volumes of text data, known as corpora. It covers the challenges of text processing and the tools used to overcome them.

- Objectives:**
1. Understand the challenges of text processing.
  2. Identify the tools and techniques used in this field.
  3. Discuss the role of corpora in language data science.
  4. Explore the impact of text processing on research and applications.

### 3.8 Project management

This learning block provides a comprehensive overview of project management, from planning and organization to execution and evaluation. It includes practical tips and exercises to help you manage your projects effectively.

- Objectives:**
1. Plan and organize a project effectively.
  2. Monitor and control project progress.
  3. Evaluate project outcomes and learn from the experience.
  4. Discuss the importance of project management in language data science.

### 3.9 Upskilling your "Intro to Language Variation" course

This learning block offers additional resources and activities to enhance your understanding of language variation. It includes links to relevant research and practical exercises.

- Objectives:**
1. Explore additional resources on language variation.
  2. Apply your knowledge to new contexts.
  3. Discuss the implications of language variation in different contexts.
  4. Explore the role of language variation in research and applications.

### 3.10 The essence of machine learning for linguists in tech

This learning block provides an overview of machine learning applications in linguistics. It covers the basics of machine learning and its role in solving linguistic problems.

- Objectives:**
1. Understand the basics of machine learning.
  2. Identify the applications of machine learning in linguistics.
  3. Discuss the challenges and solutions in this field.
  4. Explore the role of machine learning in language data science.

### 3.11 Upskilling your "Intro to Language Variation" course

This learning block offers additional resources and activities to enhance your understanding of language variation. It includes links to relevant research and practical exercises.

- Objectives:**
1. Explore additional resources on language variation.
  2. Apply your knowledge to new contexts.
  3. Discuss the implications of language variation in different contexts.
  4. Explore the role of language variation in research and applications.

### 3.12 Off the shelf games and UPSELLS

This learning block explores the use of off-the-shelf games and UPSELLS in language data science. It discusses the benefits and challenges of using these resources.

- Objectives:**
1. Understand the benefits and challenges of using off-the-shelf games and UPSELLS.
  2. Identify the applications of these resources in language data science.
  3. Discuss the role of these resources in research and applications.
  4. Explore the impact of these resources on the field.

### 3.13 Start programming with Python in 10 steps

This learning block provides a step-by-step guide to learning Python programming. It covers the basics of Python and its application in language data science.

- Objectives:**
1. Learn the basics of Python programming.
  2. Apply Python to solve linguistic problems.
  3. Discuss the role of Python in language data science.
  4. Explore the impact of Python on research and applications.

### 3.14 Why games are the way to go...

This learning block discusses the benefits of using games in language data science. It covers the role of games in research and applications.

- Objectives:**
1. Understand the benefits of using games in language data science.
  2. Identify the applications of games in this field.
  3. Discuss the role of games in research and applications.
  4. Explore the impact of games on the field.

### 3.15 D. Educational games

This learning block provides an overview of educational games and their role in language data science. It covers the design and development of educational games.

- Objectives:**
1. Understand the role of educational games in language data science.
  2. Identify the challenges and solutions in this field.
  3. Discuss the impact of educational games on research and applications.
  4. Explore the role of educational games in the future.

### 3.16 Why games are the way to go...

This learning block discusses the benefits of using games in language data science. It covers the role of games in research and applications.

- Objectives:**
1. Understand the benefits of using games in language data science.
  2. Identify the applications of games in this field.
  3. Discuss the role of games in research and applications.
  4. Explore the impact of games on the field.

### 3.17 Why games are the way to go...

This learning block discusses the benefits of using games in language data science. It covers the role of games in research and applications.

- Objectives:**
1. Understand the benefits of using games in language data science.
  2. Identify the applications of games in this field.
  3. Discuss the role of games in research and applications.
  4. Explore the impact of games on the field.

### 3.18 Why games are the way to go...

This learning block discusses the benefits of using games in language data science. It covers the role of games in research and applications.

- Objectives:**
1. Understand the benefits of using games in language data science.
  2. Identify the applications of games in this field.
  3. Discuss the role of games in research and applications.
  4. Explore the impact of games on the field.

### 3.19 Why games are the way to go...

This learning block discusses the benefits of using games in language data science. It covers the role of games in research and applications.

- Objectives:**
1. Understand the benefits of using games in language data science.
  2. Identify the applications of games in this field.
  3. Discuss the role of games in research and applications.
  4. Explore the impact of games on the field.



### 2.3 TopLang: An UPSKILLS simulation experience

Someone progresses in TopLang

- 1. Identifies learning, challenge and play path
- 2. Player proceeds to complete the next task according to the current task and can track their own progress in real time

Saving the player's choice



### 2.3 TopLang: An UPSKILLS simulation experience



### 2.3 TopLang: An UPSKILLS simulation experience



### 2.3 TopLang: An UPSKILLS simulation experience

Someone progresses in TopLang

- 1. Identifies challenge, the task path
  - 2. Through TopLang, the player gets a mission, with its goal, how to complete it and its difficulty level
  - 3. UPSKILLS tool
  - 4. Increasing level of complexity
  - 5. Feedback on wrong answers
  - 6. Areas with correct answer to collect
- ...Applying knowledge to real life problems.



### 2.3 TopLang: An UPSKILLS simulation experience

Someone progresses in TopLang

- 1. Learning path includes:
    - 2. Completing a number of tasks to allow the player with a low level of skill
    - 3. Adapting to complex tasks
      - 4. Creating a path to success
  - 2. Player is eventually asked to apply for a permanent job.
- Finally, if they choose to pursue it, they are asked to write a motivation paragraph about what they've learnt in the job (→ subject feedback)

### 3. Oh, the "fun" is over... How do I do this?

- 1. IOP: Use and learn from content collaboration from users based on feedback
  - 2. Ability to create a mix of interactive content types
  - 3. No need of user manual software for users to access levels directly in the browser
  - 4. Cross-system compatibility
  - 5. Direct integration with major LMS and even simple websites
  - 6. Easy of deployment
  - 7. "Intuitive" software
  - 8. There is a learning curve, but it has your user base in mind
- Why've got your content? Check out manuals for adapting your games available now!

### 4. And what did our students think?

- Source: Survey of student perceptions after 200hrs of TopLang
- 1. Will design and game projects be a part of my impact on student engagement and retention?
  - 2. Games used for learning should align with learner needs and enhance the learning experience.



### That's all folks!

- UPSILLUP main expected impact:
- 1. Increase prepare students for the reality of the job market
  - 2. Provide the students with respect to what skills employers are looking for
  - 3. Raise awareness among employers about the skills and aptitude of graduates of TopLang and long-term related degree
  - 4. Create engaging on-line learning content that will be freely accessible to employers
  - 5. Increase active "real" and research-based learning paths for use of educational games

## Thank you for your attention!!

Don't forget to:

1. Visit our website: <https://upskillsimulators.ai/>
2. Check out our mission: <https://upskillsimulators.ai/about/>
3. Download our reports from Zenodo: <https://zenodo.org/communities/upskills>



# Annex E: UPSKILLS learning content flyers



**SKILLS**





ADAM MICKIEWICZ  
UNIVERSITÀ DI BOLOGNA



L-Università  
ta' Malta

**The UPSKILLS Learning Content**

## A glimpse into language data science

A learning block dedicated to language data science intended as statistical analysis of data about language. It covers foundational statistical concepts involved in data description and visualisation, and it introduces statistical hypothesis testing and inference.



The UPSKILLS learning content is open access, adaptable and modular, which means that you can either wholly incorporate it in your curriculum or pick and choose any of the following modules to upskill your existing course:

- **Welcome to statistics** – populations and samples, variables, measurement (1 ECTS)
- **Working with R** – installing R, functions, objects, data preparation (1 ECTS)
- **Calculating summary numbers** – frequencies, mean, median, measures of variability (1 ECTS)
- **Showing data on graphs** – barplots, boxplots, scatterplots, mosaic plots (1 ECTS)
- **The logic behind inferential statistics** – hypothesis testing, statistical significance, inference (1 ECTS)
- **Some simple statistical tests** – Chi-square, correlation, t-tests, Wilcoxon tests (1 ECTS)
- **Student project** – putting the pieces together by studying the frequency and familiarity of words (1 ECTS)

6 ECTS  
+ a student project amounting to 1 extra ECTS

Images by storozek on Freepik



<https://upskillsproject.eu/>

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





**SKILLS**





UNIVERSITY OF  
BELGRADE

**The UPSKILLS Learning Content**

## Analytical thinking and problem solving

This learning block is focused on analytical thinking and problem solving – the soft skills that help in identifying and solving complex problems. In this block, we provide a curated selection of the most useful and user-friendly online resources, while also presenting a set of exercises in analytical thinking and problem solving designed explicitly for language and linguistics specialists.



The UPSKILLS learning content is open access, adaptable and modular, which means that you can either wholly incorporate it in your curriculum or pick and choose any of the following modules to upskill your existing course:

- **What is it all about?** – basic concepts and methods (1 ECTS)
- **Practice makes perfect** – exercises in analytical thinking and problem solving for language and linguistics specialists (1 ECTS)
- **Think, solve, repeat** – What do you learn when you study language and linguistics (1 ECTS)
- **Student project** – completing the learning experience (1 ECTS)

3 ECTS  
+ a student project amounting to 1 extra ECTS

Images by storozek on Freepik



<https://upskillsproject.eu/>

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





**SKILLS**





CLARIN

**The UPSKILLS Learning Content**

## Automatic speech recognition and forced alignment

A learning block aimed to make a bridge between students with a linguistic background and Automatic Speech Recognition as a technical topic. It is set up to present a transparent and balanced presentation of ideas, at appropriate levels of detail, alongside quizzes containing multiple-choice questions which lecturers are invited to adapt and expand as research in the field progresses.



The UPSKILLS learning content is open access, adaptable and modular, which means that you can either wholly incorporate it in your curriculum or pick and choose any of the following modules to upskill your existing course:

- **The speech signal**
- **Acoustic features**
- **Bayes and Viterbi**
- **Architectures of ASR (I)**
- **Architectures of ASR (II)**
- **Forced alignment as a special case of ASR**
- **Data selection criteria / justification**
- **Dialogue**
- **Language models**
- **Student project**

3 ECTS  
+ a student project amounting to 3 extra ECTS

Images by storozek on Freepik



<https://upskillsproject.eu/>

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





**SKILLS**





UNIRI



UNIVERSITÄT GRAZ  
UNIVERSITY OF GRAZ

**The UPSKILLS Learning Content**

## Collecting language data from human participants

A learning block dedicated to human language data collection. Structured into four modules, it addresses methods for collecting language data from human participants, with a focus on judgement and elicited production tasks, experimental tasks to test comprehension and production in the second language, as well as the conduct of ethnographic fieldwork in sociolinguistics.



The UPSKILLS learning content is open access, adaptable and modular, which means that you can either wholly incorporate it in your curriculum or pick and choose any of the following modules to upskill your existing course:

- **General** (1 ECTS): Introduction to linguistic research with human participants, ethics in linguistic research with human participants, population sampling in linguistic research projects
- **Word-level phenomena** (1.5 ECTS): Word-level phenomena: Judgement and production tasks
- **Second language acquisition** (0.75 ECTS): Comprehension, production and acceptability judgement tasks in second language acquisition
- **Sociolinguistics** (1.75 ECTS): Ethnographic fieldwork, surveys and questionnaires in sociolinguistics, the sociolinguistic interview.
- **Student projects**: "Irregular" phonological alternations (3 ECTS) & Second language acquisition of English morphosyntax (3 ECTS)

6 ECTS  
+ two student projects amounting to 6 extra ECTS

Images by storozek on Freepik



<https://upskillsproject.eu/>

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



**UP SKILLS**

UNIVERSITY OF BELGRADE  
ALMA MATER STUDIORUM UNIVERSITÄT DE BOLOGNA

### The UPSKILLS Learning Content

## First steps into scientific research




The UPSKILLS learning content is open access, adaptable and modular, which means that you can either wholly incorporate it in your curriculum or pick and choose any of the following modules to upskill your existing course:

- **Introducing science** – defining what science and scientific research are about (1 ECTS)
- **The big picture** – approaches to research, research questions and hypotheses (1 ECTS)
- **A look under the hood** – variables and research design (1 ECTS)
- **Student project** – putting the pieces together by studying language and gender (1 ECTS)

3 ECTS  
+ a student project amounting to 1 extra ECTS

<https://upskillsproject.eu/>

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

**UP SKILLS**

CLARIN

### The UPSKILLS Learning Content

## Introduction to language data: Standards and repositories




The aim of this learning block is to provide lecturers in BA/MA language-related programmes with a pool of learning resources and activities that they can use in the classroom to introduce students to research data repositories and their role in the linguistic research data lifecycle in the context of Open Science and FAIR data principles.

6 ECTS  
+ a student project amounting to 1 or 2 extra ECTS

The UPSKILLS learning content is open access, adaptable and modular, which means that you can either wholly incorporate it in your curriculum or pick and choose any of the following modules to upskill your existing course:

- **Introduction to the language resource lifecycle and management**
- **How research data repositories help make language data FAIR**
- **Finding and (re)using language resources in the CLARIN repositories**
- **Citing language and linguistic data**
- **Legal and ethical issues in language data collection, sharing and archiving**
- **Student project** – designing, compiling and archiving a corpus of bank bulletins
- **Glossary**

<https://upskillsproject.eu/>

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

**UP SKILLS**

UNIVERSITY OF BELGRADE  
ALMA MATER STUDIORUM UNIVERSITÄT DE BOLOGNA

### The UPSKILLS Learning Content

## Processing texts and corpora




A learning block that puts together basic notions and techniques of text processing with skills related to corpus linguistics, such as corpus creation and management and the ability to use corpus query tools for simple and advanced queries. Characterized by a largely practical and research-based approach, this block links theoretical and technical skills to real-life working situations, while aiming to generate curiosity, interest and nurture critical thinking in the minds of the students.

The UPSKILLS learning content is open access, adaptable and modular, which means that you can either wholly incorporate it in your curriculum or pick and choose any of the following modules to upskill your existing course:

- **Why process texts?**
- **Basics of text processing**
- **Corpus design and construction**
- **Corpus annotation**
- **Corpus consultation**
- **Corpus types, research priorities and applications**
- **Student project** – speaking of consequences: good, bad, neutral?

5 ECTS  
+ a student project amounting to 1.5 extra ECTS

<https://upskillsproject.eu/>

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

**UP SKILLS**

UNIVERSITY OF BELGRADE

### The UPSKILLS Learning Content

## Project management




This learning block is focused on the knowledge and tools needed to achieve specific goals within specific constraints. In it, we provide a curated selection of the most useful and user-friendly online resources. In addition, we present a set of exercises in project management specifically designed for language and linguistics specialists. Finally, we provide a student project task to complete the learning experience.

3 ECTS  
+ a student project amounting to 1 extra ECTS

The UPSKILLS learning content is open access, adaptable and modular, which means that you can either wholly incorporate it in your curriculum or pick and choose any of the following modules to upskill your existing course:

- **The nuts & bolts** – the notion and characteristics of a project; general concepts of project management and professional standards; the profile of a project manager; main project phases (i.e. project management life cycle); project scheduling and implementation. (1.5 ECTS)
- **Getting organized** – exercises in project management for language and linguistics specialists. (1.5 ECTS)
- **Student project** – getting it done: Managing translation projects (1 ECTS)

<https://upskillsproject.eu/>

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

**UP SKILLS**

University of Zurich  
L-Università ta' Malta

**The UPSKILLS Learning Content**

**Start programming with Python in 10 steps**



The UPSKILLS learning content is open access, adaptable and modular, which means that you can either wholly incorporate it in your curriculum or pick and choose any of the following modules to upskill your existing course.

- Enter the world of programming
- Watch Python code demo
- Write and run your first programs
- Figure out some details
- Get out of your notebooks
- Start matching with regular expressions
- Grow your dictionaries
- Web is the link
- Organise your code with functions and classes
- Contribute your code

5 ECTS

<https://upskillsproject.eu/>

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

**UP SKILLS**

University of Zurich  
L-Università ta' Malta

**The UPSKILLS Learning Content**

**The essence of machine learning for linguists in tech**



The UPSKILLS learning content is open access, adaptable and modular, which means that you can either wholly incorporate it in your curriculum or pick and choose any of the following modules to upskill your existing course.

- Machine learning needs linguists
- Things in mathematical space
- Data points need labels
- Setting boundaries with the perceptron algorithm
- Linguists need neural networks
- Meanings are vectors
- Learning meanings with (large) language models
- NLP makes for happy users
- How good is an NLP model?
- The practice and ethics of large language models (LLMs)

5 ECTS

<https://upskillsproject.eu/>

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

**UP SKILLS**

UNIVERSITÉ DE GENÈVE

**The UPSKILLS Learning Content**

**Upskilling your "Introduction to Language Variation" course**



This learning block has a double function: it can be considered as a concrete example of how to teach a pre-existing 'Introduction to Language Typology / Language Variation' course while adopting a "game-based" learning approach, or it can be imported as it is and used as teaching material for an 'Introduction to Language Variation' course.

The UPSKILLS learning content is open access, adaptable and modular, which means that you can either wholly incorporate it in your curriculum or pick and choose any of the following modules to upskill your existing course:

- Introduction to language variation (0 ECTS – compulsory unit)
- Linguistic data collection (2 ECTS)
- Theory-driven data analysis (2 ECTS)
- Student project – Researching classifiers cross-linguistically (2 ECTS)

4 ECTS  
+ a student project amounting to 2 extra ECTS

<https://upskillsproject.eu/>

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