Adapting UPSKILLS Learning Modules to the University Curricula: Best Practices and Lessons Learnt from the H2IOSC Training Experience at the University of Ferrara

Giulia Pedonese
CNR Institute of Computational Linguistics "Antonio Zampolli", Pisa, Italy
giulia.pedonese@cnr.it

Francesca Frontini
CNR Institute of Computational Linguistics "Antonio Zampolli", Pisa, Italy
francesca.frontini@cnr.it

Dario Del Fante
Department of Humanities University of Ferrara, Ferrara, Italy
dario.delfante@unife.it

Eleonora Federici
Department of Humanities University of Ferrara, Ferrara, Italy
eleonora.federici@unife.it

Abstract

This paper details the steps taken to adapt and integrate the training materials developed by CLARIN ERIC in two bachelor's degree courses and one master's degree course at the University of Ferrara. The workflow applies the shared methodology developed within the Humanities and Heritage Italian Open Science Cloud project. It modifies the training materials of the UPSKILLS course "Introduction to Language Data: Standards and Repositories" according to the needs of three target courses focusing on English to Italian translation: English Language Course for Tourism, English Language for Translation and English Language and Linguistics for Humanities, Arts and Archaeology. The result of this pilot is a documented example of how CLARIN services can be integrated into university teaching, including initial teacher training, and providing an opportunity to discuss the topic and a use case for trainers who intend to include CLARIN in their courses.

1. Introduction

The Humanities and Cultural Heritage Italian Open Science Cloud project (H2IOSC) (Degl'Innocenti et al. 2023) aims to create a federated cluster of the services and resources developed by the national nodes of four research infrastructures for Open Science that are part of the European Strategy Forum on Research Infrastructure (ESFRI) roadmap in the area of social and cultural innovation.[1] One is CLARIN-IT, the Italian consortium of the Common Language Resource and Technology Infrastructure.[2] In line with other projects of national and international scope, H2IOSC devotes an entire work package to training and education: the Work Package 8 (WP8), whose aim is to define a comprehensive shared strategy for training at the level of single infrastructures and for the whole cluster based on the needs of

[1] H2IOSC Project - Humanities and cultural Heritage Italian Open Science Cloud funded by the European Union NextGenerationEU - National Recovery and Resilience Plan (NRRP) - Mission 4 "Education and Research" Component 2 "From research to business" Investment 3.1 "Fund for the realization of an integrated system of research and innovation infrastructures" Action 3.1.1 "Creation of new research infrastructures strengthening of existing ones and their networking for Scientific Excellence under Horizon Europe" - Project code IR0000029 - CUP B63C22000730005. Implementing Entity CNR.

[2] The participating research infrastructures are: Digital Research Infrastructures for the Arts and Humanities (DARIAH); European Research Infrastructure for Heritage Science (E-RIHS), Common Language Resource and Technology Infrastructure (CLARIN) and Open Scholarly Communication in the European Research Area for Social Sciences and Humanities (OPERAS). See the official website: https://www.h2iosc.cnr.it/home/

the community as identified with the landscaping activity of Work Package 2. The repurposing of reusable teaching materials planned within H2IOSC WP8 is in line with the teaching strategy of other European projects: first, with the UPSKILLS[3] project, which aims to integrate university curricula of language subjects with skills required by the labor market, and for which CLARIN developed the course "Introduction to Language Data: Standards and Repositories" (Section 2). Secondly, repurposing teaching materials for different audiences aligns with the aims of the Skills for the European Open Science Commons project (Skills4EOSC)[4] to expand European researchers' skills by creating European competence centres, with which H2IOSC training has aligned (Section 2). Therefore, this initiative appears to be embedded in a broader context, which aims not only to offer educational materials suited to the professional needs of a learner audience but also to create a network of trainers who can exchange skills and materials.

This paper aims to show the application of the training strategy detailed in WP8 to the modules designed by CLARIN-IT within activity 8.2: "Teach CLARIN, teach with CLARIN: training, communication and impact", adapting them to the training needs of the University of Ferrara. This experience resulted in a case study to show the potential of such a strategy when integrating the CLARIN research infrastructure into teaching. First, there will be an overview of the H2IOSC training methodology and its applications to CLARIN courses (Section 2) and a background on the training needs of the selected curricula at the University of Ferrara (Section 3). Then, Section 4 and its subsections will detail the steps of this pilot. Finally, Sections 5 and 6 will highlight the steps taken in 2025 and the lessons learnt from this experience.

2. H2IOSC Training Methodology Applied to CLARIN Courses

The H2IOSC project's training strategy is aimed primarily at the Italian Social Sciences and Humanities communities and offers tailored training on the resources and tools available in the participating infrastructures. An essential pillar of this strategy is the train-the-trainers perspective, which aims to help teachers of university and professional courses enhance their skills and those of their students. To this end, in WP8, we developed a shared methodology for designing training materials as digital objects compliant with the FAIR principles (Wilkinson et al., 2016). We implemented two training platforms to facilitate their reuse: a Learning Management System for students to access course materials in a modular and intuitive way, and a repository for teachers to browse and download learning resources from. As the development of those platforms was still ongoing while the lessons in Ferrara took place, we will focus on the methodology we applied (Pedonese et al. 2024) to make the course materials accessible and reusable on different devices. However, the integration of these technologies is planned in the continuation of the pilot.

While developing a methodology to align all four research infrastructures on the adaptation and creation of reusable training materials, we identified the *FAIR-by-Design Methodology for Learning Materials* as the standard established by the Skills4EOSC project, funded by the European Union under Horizon Europe and aimed to enhance Open Science skills in Europe by promoting FAIR practices. This framework ensures training materials align with the FAIR principles (Findable, Accessible, Interoperable, Reusable) to enhance their reusability in the scientific community. It adopts a backwards instructional design process, expanded with six steps: planning, development, quality check, publication, and dissemination, emphasising metadata, interoperability, and granularity and includes practical tools like checklists and templates to support implementation (Filiposka et al. 2023). The primary references we have drawn from the methodology, adapting it to the needs of our disciplinary area, have been the adoption of a standard metadata schema and the definition of the learning object as the smallest unit to which it applies.

[3] UPSKILLS stands for "UPgrading the SKIlls of Linguistics and Language Students": it was an Erasmus+ strategic partnership for higher education running from September 2020 to August 2023. See the official website: https://upskillsproject.eu/

[4] It ran from 2022 to 2025, with the goal of building a comprehensive training ecosystem. The project brought together international partners to advance open, interoperable scientific practices. See the official website: https://www.skills4eosc.eu/

To effectively describe and share H2IOSC learning materials, we adopted the metadata set developed by the Research Data Alliance (RDA), the Minimal Metadata Schema for Learning Resources (Hoebelheinrich et al. 2022). As shown in Fig.1, this dataset includes 14 fields divided between descriptive information (title, author, etc.), access information (licence, URL, etc.) and didactic information (type of resource, target group, initial level, learning outcomes).

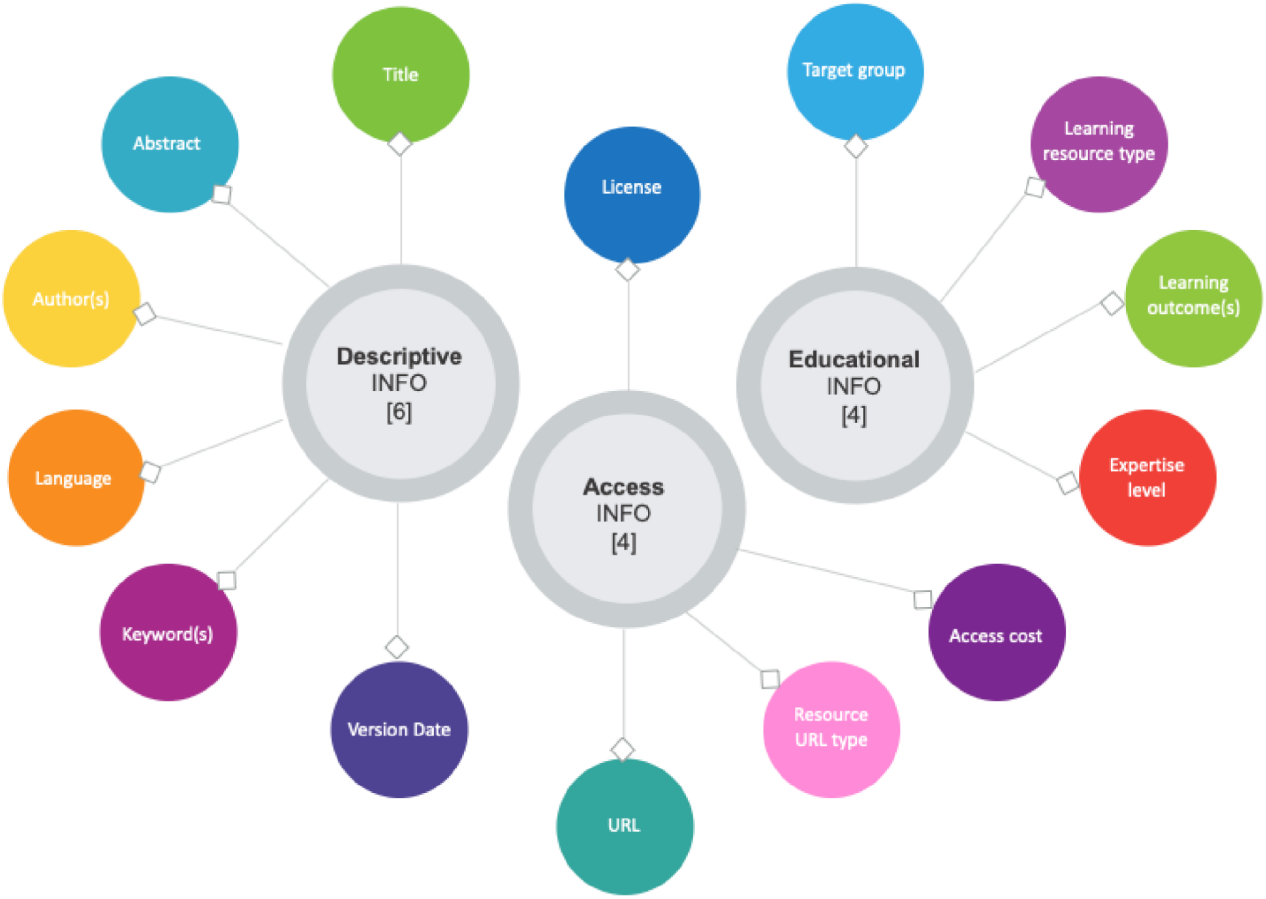

Fig. 1. RDA Minimal Metadata Set with Definitions (Hoebelheinrich et al. 2022).

This highly flexible model has already been adapted to the needs of CLARIN (van der Lek et al. 2023 c), which has invested in training by promoting standard practices and creating a community of trainers through the CLARIN Trainers' Network. Fig. 2 provides an example of the CLARIN adaptation. In H2IOSC, we have adopted the modified model by including specific fields such as *contributors*, *workload in ECTS*, *PID*, *version date,* and *standard citation*.

Metadata to describe the CLARIN Training Materials

Title	The title of the training material.
Abstract/Description	Describe the topic, general goals and objectives of the training materials.
Author (s)	Name of entity (ies) authoring the materials.
Contributor (s)	Name of entity (ies) contributing to the development of the training materials.
(Sub)discipline (s) & Topic	Indicate the (sub)discipline or cluster & the topic (e.g. social sciences/research data management).
Training Material Type	Indicate the type of training material (e.g. presentation, tutorial, e-learning module, course, unit/lesson, report, video, webinar, slides, game).
Primary Language	Indicate the language (s) in which the materials were originally published or made available. If the training material is in a language other than English, please include an English summary in the ReadMe file.
Keywords	Keywords describing the training materials to improve search and discoverability.
Workload (in ECTS, if applicable)	Describe the structure of the materials and the settings in which to deliver them, including the time allocated to each part (lectures, exercises, etc.).
URL to Training Material	URL that resolves to the training materials or to a "landing page" for the materials that contains contextual information, including the direct resolvable link to the training materials, if applicable.
Persistent Identifier	The identifier assigned to the materials, e.g. DOI, Handle, ARK.
Target Audience	Principal users for which the training material was designed.
Target Skills Level	Target skill level in the topic being taught (e.g. beginner, intermediate, advanced).
Training Material Type	Indicate the type of training material (e.g. presentation, tutorial, e-learning module, course, unit/lesson, report, video, webinar, slides, game).
Learning Outcomes	Describe what knowledge, skills and abilities a learner should acquire upon completing the training/course. Please use Bloom's Taxonomy to describe the outcomes.
CLARIN Resources Used in the Training	Cite the CLARIN resources, NLP tools, repositories and other services used in the training/course.
Facilities Required	Technical resources and related materials (software requirements, access to specific datasets or infrastructure services)
Licensing and (Re)use Details	The licence under which the materials are shared, and rules and conditions for (re) use and contribution.
Preferred Citation	Instructions on how to cite your material.
Creation Date and Last Revision	Indicate the creation and last modification date of the training material.
Version Number, if applicable	Version date for the most recently published material.

Fig. 2. Metadata to Describe the CLARIN Training Materials (van der Lek et al. 2023 c).

According to the best practices recommended in Skills4EOSC, we considered the single lesson as a learning object, the minimum unit to apply the FAIR principles and be described with the RDA metadata schema. The learning object, as defined in the methodology, is "any digital resource that supports learning developed around a single learning objective defined as a package of a lesson, activity and assessment with a concrete learning outcome" (Filiposka et al. 2023:8).[5] While this certainly facilitates the reusability of training materials for future reuse, during the course delivery phase to students, it was applied to teaching by dividing the module into several lectures of no more than 25 minutes alternating with multiple-choice quizzes using interactive software such as *Kahoot!* and *Mentimeter*, implementing gamification strategies. Furthermore, we adopted open Creative Commons licenses and standard citations while sharing the materials with students and improved the accessibility of teaching materials by converting source formats into even more open and editable formats, possibly open source.[6]

In the first half of the project, the CLARIN-IT consortium, represented by its founder member and host of the ILC4CLARIN service provider centre at the CNR Institute of Computational Linguistics,[7] worked on applying the Skills4EOSC methodology to the UPSKILLS course "Introduction to Language Data: Standards and Repositories" (van der Lek et al. 2023 a) which was already reusable in compliance

[5] For a more extensive definition, see also pp. 23-24 of the deliverable.
[6] More detais on the H2IOSC methodology for training materials in Pedonese et al., 2024 now submitted for reviewing in Umanistica Digitale.
[7] https://ilc4clarin.ilc.cnr.it/

with the FAIR principles and accessible on the project's Moodle platform.[8] As shown in Fig.3, the user could download each module of this course by selecting it from the course structure.

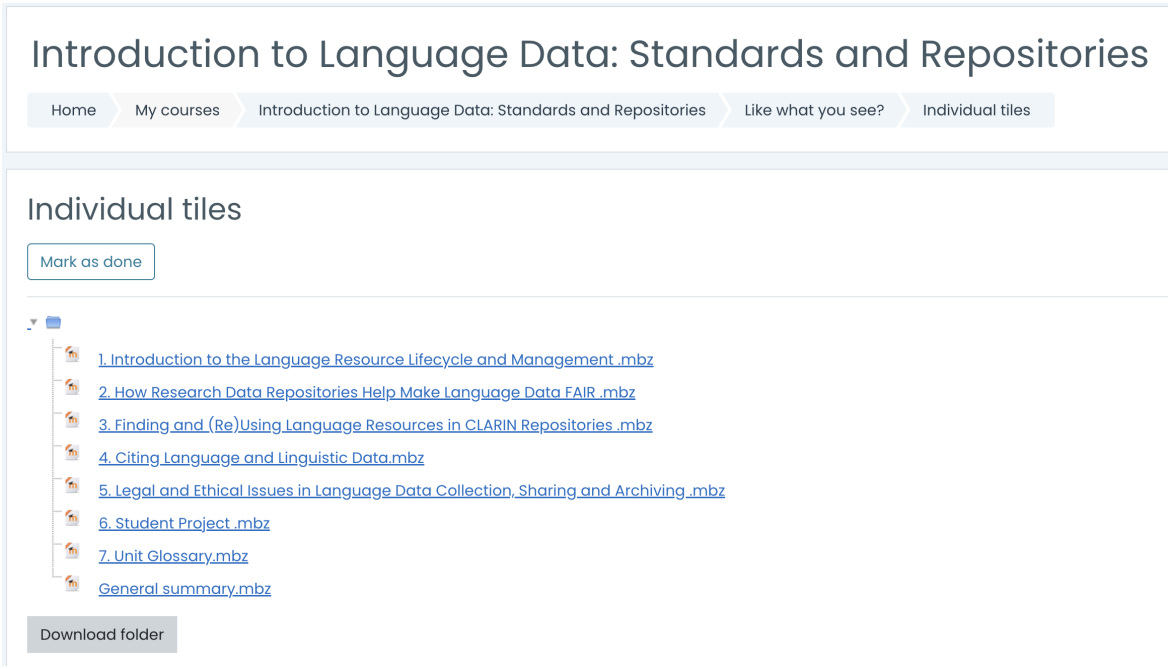

Fig. 3. Original Training Materials on UPSKILLS Moodle Platform (van der Lek et al., 2023 a).

The course is designed to provide instructors and students in BA/MA linguistic programs with educational resources and practical activities on using research data repositories within the lifecycle of linguistic data, aligning with the FAIR and Open Science principles. Its learning objectives are using certified repositories to discover, share, publish, and store linguistic resources and datasets and applying integrated services and tools from repositories to process, annotate, and analyse various types of corpora according to community standards.

To maximise its overall reusability and adapt the course materials ourselves, we first asked the creators to access the editable version of the learning block, since the only available version was in Moodle predefinite format (.mbz file), which we could download from the original project's Moodle and was only editable via the same platform. We then translated the entire course into Italian and published it as a structured aggregate of Markdown[9] files following the Skills4EOSC methodology (van der Lek et al. 2024). Fig. 4 shows an example of training material in a Markdown file: this highly flexible and compatible annotation allows users to easily integrate these training materials into their editing tools.

[8] https://upskillsproject.eu/. The course is available on the UPSKILLS Moodle platform: https://upskillsproject.eu/project/standards_repositories/. For CLARIN's activities in UPSKILLS and the training materials produced see also https://www.clarin.eu/content/upskills-learning-and-teaching-materials (Gledić et al. 2023).

[9] https://www.markdownguide.org/

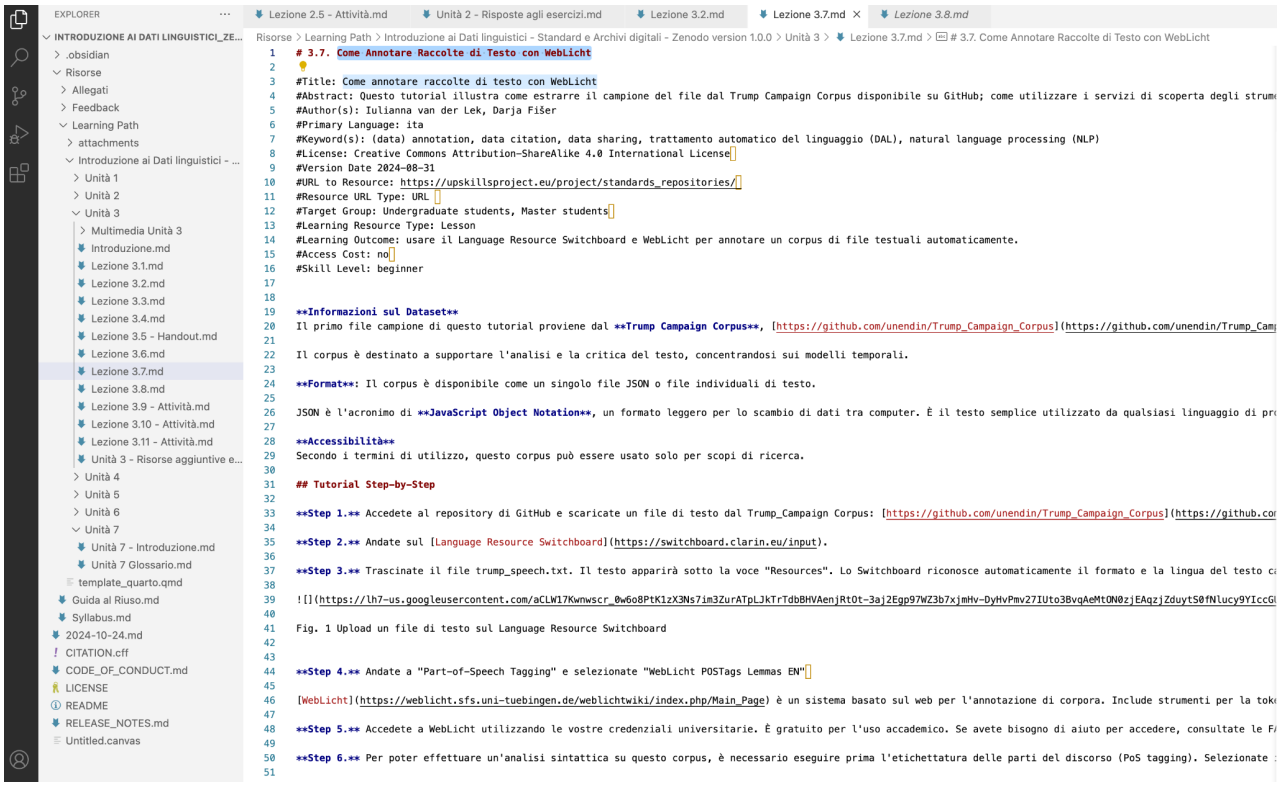

Fig. 4. Training Materials' Adapted Structure in Markdown Editor (van der Lek et al., 2024).

While the course was being translated, we selected the parts that could be reused for teaching at the University of Ferrara: we adapted Lessons 1.1 to 1.4 on learning resources, disciplinary repositories and lessons 3.1 to 3.3. on CLARIN core services that implement the FAIR principles, such as the Virtual Language Observatory and the Language Resource Switchboard, and an introduction to the fundamentals of Natural Language Processing. Those were the main points on which we wanted to raise the interest of the students of the selected courses in Ferrara (Section 3), as language studies, both in academia and business, can no longer do without proper data management.

3. Teaching Digital Humanities at the University of Ferrara

The University of Ferrara is a dynamic environment where digital humanities have the potential to fruitfully interact with linguistic research, especially with CLARIN's support. The University Language Centre of the University of Ferrara has been a CLARIN-IT member since 2023, providing metadata related to language resources for the ILC4CLARIN repository centre. Thus, integrating CLARIN services into teaching is a way to educate future researchers and professionals in an increasingly digitised context, providing them with tools and resources to support them in their careers. This is why a collaboration was established between the CNR-ILC personnel dedicated to training in the H2IOSC project and the CLARIN-IT referent for the University Language Centre, who also happen to be lecturers of linguistic courses at BA, MA and PhD level.

More specifically, the collaboration aimed to acquaint the students with the fundamentals of digital humanities using CLARIN tools and services. To this end, we selected three courses Prof. Federici and Dr. Del Fante taught. We included two lessons at the end of each course, providing an introduction to the use of language resources and language technologies, how to access CLARIN tools, how to use a repository such as ILC4CLARIN, and why practices for accessing, publishing and reusing data on the web like the FAIR principles can have a direct impact on their student careers. So, we chose three English Language and Linguistics courses belonging to three different degree programs respectively: 1) Bachelor's degree in Humanities, Arts, and Archaeology; 2) Master's degree in Foreign Languages and Literature; 3) Bachelor's degree program in Manager of Cultural Itineraries.

The English Language for Humanities, Arts, and Archaeology course (1) aims to provide students with critical tools to understand English information. The course focuses on three main topics: English

critically approached as a global language, the notion of discourse, and communicative strategies for representing people and identities in political communication and social media contexts. The English Language course for Modern Languages and Literature (2) consists of a theoretical part, which offers a panorama of the main aspects of Cultural Translation, and a practice section, which will translate literary texts of various periods and contexts from English to Italian. Finally, the English Language for Tourism course (3) aims to equip students with linguistic and metalinguistic tools essential for developing an awareness of the English tourist language. This goal is achieved through linguistic and cultural analysis of various contemporary tourism text types. These were selected as pedagogical spaces to introduce students to CLARIN technologies, and the UPSKILLS course was adapted to fit each skill level and focus. The adaptation process was relatively easy, as the course materials were already structured in a highly flexible way, allowing us to combine multiple topics and adjust the content as needed.

As this was the first experiment of this kind, we tried to include different disciplinary areas encompassing a broad selection of professional profiles, to test the flexibility of the teaching materials that were made available. In parallel, we devised a method to gather as much feedback as possible to improve the teaching-learning experience for future editions.

4. CLARIN-IT Training at the University of Ferrara

The training initiative stemmed from preexisting contacts between the CNR-ILC and the University of Ferrara and followed the train-the-trainer approach. The first meeting was held online to align Dr. Dario Del Fante, teacher and researcher at the University of Ferrara, with the latest available resources, define the lesson plan, and schedule an assessment strategy. Then, we held the lessons in person at the University of Ferrara between April and May 2024, finally assessed the students, and gathered feedback via a questionnaire. The following subsections will detail the phases of the pilot as shown in Fig. 5.

Fig. 5. Phases of the Pilot.

4.1 Training the Trainer

The first meeting was held online and consisted of the presentation of the primary teaching resources currently available for the Italian SSH community. The materials from the adapted UPSKILLS course and the facilitation guide developed within the H2IOSC project on how to reuse them were presented. A precious resource in this phase was the UPSKILLS "Guidelines on integrating Research Infrastructures into Teaching: Recommendations and Best Practices" (van der Lek et al. 2023 b), an up-to-date tool to enhance research-based teaching, presenting the case study of CLARIN. Since Dr. Del Fante was already familiar with CLARIN core services, the training session focused mainly on their latest functionalities and the discovery process of language resources that could be relevant to students.

Upon requests for specific translation technologies to teach students, further inquiries were made by the CLARIN-IT trainer, which resulted in a fruitful discussion with the CLARIN ERIC Training Officer, Iulianna van der Lek. This led to the definition of a CLARIN Café on Translation Technologies and Workflows for SSH Research,[10] held on June 14, 2024. The seminar was aimed at providing instructors, researchers, and CLARIN infrastructure staff with the tools to introduce these technologies in their teaching and research activities.[11] As the initiative originated directly from the training needs of CLARIN consortium members and was then extended to the infrastructure's broader community via the CLARIN Trainers' Network, it was one of the many results of the collaboration with the University of Ferrara.

[10] https://www.clarin.eu/event/2024/clarin-cafe-translation-technologies-workflows-ssh-researchers
[11] https://www.h2iosc.cnr.it/h2iosc-training-clarin-cafe/

4.2 Training the Students

For teaching the students of all three courses, we developed three modules of two lessons each to be held in person in Ferrara: 1) English Language Course for Tourism: two lessons of two hours each for 40 students; 2) English Language for Translation: two lessons of two hours each for 25 students; 3) English Language and Linguistics for Humanities, Arts and Archaeology: two lessons of two hours each for 30 students. For clarity, the lesson plan is shown in Fig. 6.

Course	Level	Lesson 1	Lesson 2	Total duration	Attendance
English Language for Translation	MA	April 18 Teacher: Dr. Del Fante	April 19 Teacher: Dr. Del Fante	4 hours	25 students
English Language and Linguistics for Humanities, Arts and Archaeology	BA	May 9 Teacher: Dr. Pedonese	May 17 Teacher: Dr. Del Fante	4 hours	30 students
English Language for Tourism	BA	May 10 Teacher: Dr. Pedonese	May 14 Teacher: Dr. Del Fante	4 hours	40 students

Fig. 6. Lesson Plan.

For the first lesson of each module, we adopted the same format, tailoring the examples and case studies to the disciplinary interests of each course. This introductory lesson offered a detailed overview of the CLARIN research infrastructure, emphasising the support provided by CLARIN Knowledge Infrastructure and giving information on funding opportunities. Additionally, we introduced the use of CLARIN Language Resource Families and the main functions of the metacatalogue Virtual Language Observatory.

We then dedicated the second lesson of each module to the practical application of CLARIN core services, and, in this case, we tailored the content to the specific needs of each course. Each resource and tool we chose to show the students was accessed through the Virtual Language Observatory to familiarise the class with CLARIN services, and we encourage students of all three courses to include those digital tools in their research projects. In the first course, we implemented the recently released ParlaMint 4.0[12] to analyse cross-linguistic political communication in Italy and the UK: we demonstrated how to use metadata to analyse textual data differently and provide a more comprehensive analysis of the texts. In the second course, we focused on using the parallel corpora available in the CLARIN infrastructure to assist translators. Specifically, we utilised the Intercorp parallel corpus (Čermák & Rosen 2012) to provide exercises addressing translation issues. We also utilised *Treq - the translation equivalent database*[13] - to compare the best translation equivalents and synonyms available in the corpus. For the third course, we demonstrated the utility of Voyant Tools in analysing the most recurrent features of promotional language used on different international tourist websites.

[12] ParlaMint is a CLARIN flagship project aimed at facilitating cross-linguistic and cross cultural analysis of political discourse by developing a multilingual and comparable corpus of parliamentary debates from 29 European countries convering the period from 2015 to 2022. The corpus contains over 1 billion words and features rich metadata on approximately 24,000 speakers. The data is uniformly encoded and linguistically annotated up to the level of Universal Dependencies syntax and named entities: https://www.clarin.eu/parlamint ParlaMint 4.0, released in October 2023, expanded the coverage, enhanced metadata and provided new data quality improvements.https://lindat.mff.cuni.cz/services/teitok/parlamint-40/index.php?action=subselect&id=GB

[13] https://treq.korpus.cz/index.php

5. Feedback and Assessment

Since this is a developing project, we submitted a feedback form to the students along with the final quiz to assess their knowledge of the course's core concepts in the short term and the quality of the course itself. The feedback and assessment form was the same for all three classes, and 20 participants completed it. In the quiz, students were asked questions such as defining Open Science, choosing the correct sequence of FAIR principles and completing a Data Management Plan definition. An average of 80% of the students answered these questions correctly, showing they had learnt about those topics and could comprehend and use CLARIN core services. As for the quality feedback, we asked the students to rate their satisfaction, effort, and relevance of the course activities on a scale of 1 to 5.

An important aspect to note was student satisfaction, which scored 3.67 out of 5. This figure, cross-referenced with the good performance of the participating students in the assessment exercises, shows that the course was well-tuned to the students' abilities. On the other hand, the students' perception of the usefulness of the topics covered was a critical aspect, which scored 3.42 out of 5. Even though lecturers always emphasised the possibilities of the CLARIN tools demonstrated in lectures, including this information within the final part of the courses may have been ungraded. They may have led students to think it was a secondary part.

For this reason, we have decided to reshape the second edition of the course in the 2024-2025 academic year by devoting more time to these topics from the first semester of classes and selecting a smaller cohort of students upstream, focused exclusively on language studies. This allowed us to take a more integrative approach to the students' university curriculum, which could accompany them in developing skills geared towards professional figures leaving university and eventually employed in the publishing and cultural industries.

6. Further Developments

Starting from the lessons learnt in this first experiment, for the 2024-2025 edition, which is still ongoing, we restricted the selection to two courses explicitly focusing on linguistics and translation technologies: English Language (MA) and English Language III (BA). We have also strengthened the assessment plan by including the final quiz and feedback form as necessary elements of the course evaluation to encourage all students' participation. In addition, we decided to add a preliminary questionnaire to test students' prior knowledge of the topics covered in the lectures, which in this case focused on machine and computer-aided translation tools. Surprisingly, all the students already use machine translation tools such as DeepL and Google Translate in their everyday lives, but no students can write down a definition of these technologies. So, we have tried to bridge the gap between their knowledge and their actual use of these tools, delving into the workings of the different types of machine translation and highlighting the possibilities of integrating them into computer-assisted tools such as MateCat.[14]

In addition, this time, we adapted and reused the materials of a new H2IOSC course along with the previous year's UPSKILLS materials. In a specific lesson restricted to MA students, we illustrated the fundamentals of the Linked Data paradigm, repurposing a part of the course Linguistic Linked Open Data for Humanists, which was initially created for international MA and PhD students participating in the Lisbon Summer School in Linguistics (Khan et al. 2024). As this new version of the pilot at the University of Ferrara is still in progress, we only have partial results from the first three lessons held in November 2024. However, we already saw a significant shift in the students' participation and engagement thanks to the more structured assessment plan and the interactive teaching strategies we implemented, primarily through gamification software like *Kahoot!* and Mentimeter.

Since the collaboration between CNR-ILC and the University of Ferrara remains active, we continued the lectures in the second semester. We further adapted the approach to another English language course for students in the master's degree program in Education, Communication and Digital Citizenship. In the lecture on April 9, data management aspects were emphasised, and a Data Management Plan

[14] https://www.matecat.com/

template was proposed along with the presentation of the data steward, who could be one of the professional outlets envisaged by the course of study.

Beyond the individual adaptations, the relevant aspect is the scalability of the approach, which includes an instructional design methodology that ensures easy reuse and updating of the material. The starter course, "Introduction to Language Data: Standards and Repositories," was very intuitive to repurpose because it was structured in highly editable modules. Subsequent repurposing in H2IOSC allowed it to be placed within an integrated teaching infrastructure, and it ensured that a wider cohort of students with diverse backgrounds and different professional profiles could access the material.

7. Results

This series of training events at the University of Ferrara allowed us to collect insight into using CLARIN language resources and services in academic teaching activities and how to develop new modules reusing existing courses, namely the UPSKILLS training materials. Thanks to this experience, we could better understand the steps needed to train a trainer who is already familiar with CLARIN resources but has new and challenging requests to which CLARIN-IT needs to respond with the help of the broader international consortium. Another lesson learnt is the necessity of adapting methods, standards, tools and resources to make them relevant to the Italian community, which is only possible thanks to disseminating success stories on both national and international levels.

References

Čermák, F. – Rosen, A. 2012. The case of InterCorp is a multilingual parallel corpus. In *International Journal of Corpus Linguistics*, 17(3), 411–427.

Degl'Innocenti, Emiliano, Monica Monachini, Alberto Bucciero, Enrico Pasini, Bruno Fanini, and Francesca Frontini. 2023. 'H2IOSC: Humanities and Heritage Open Science Cloud'. In *La Memoria Digitale: Forme Del Testo e Organizzazione Della Conoscenza. Atti Del XII Convegno Annuale AIUCD*, edited by Emmanuela Carbé, Gabriele Lo Piccolo, Alessia Valenti, and Francesco Stella, 63–64. https://iris.unive.it/retrieve/0f226d38-e332-418b-9b14-d5558d1a0d9d/AIUCD2023.pdf.

Filiposka, Sonja. 2023. 'D2.2 Methodology for FAIR-by-Design Training Materials', August. https://doi.org/10.5281/ZENODO.8305540.

Frontini, Francesca, and Monica Monachini. 2023. 'Infrastrutture Digitali per Le Scienze Umane e Sociali'. In *Digital Humanities. Metodi, Strumenti, Saperi*, 197–213. Roma: Carocci. https://www.carocci.it/prodotto/digital-humanities.

Garcia, Leyla, Bérénice Batut, Melissa L. Burke, Mateusz Kuzak, Fotis Psomopoulos, Ricardo Arcila, Teresa K. Attwood, et al. 2020. 'Ten Simple Rules for Making Training Materials FAIR'. *PLOS Computational Biology* 16 (5): e1007854. https://doi.org/10.1371/journal.pcbi.1007854.

Gledić, Jelena, Maja Đukanović, Jelena Budimirović, Nađa Soldatić, Maja Miličević Petrović, Silvia Bernardini, Adriano Ferraresi, et al. 2023. 'UPSKILLS Teaching and Learning Content'. *Https://Upskillsproject.Eu/*, August. https://www.clarin.si/repository/xmlui/handle/11356/1865.

Hoebelheinrich, Nancy J, Katarzyna Biernacka, Michelle Brazas, Leyla Jael Castro, Nicola Fiore, Margareta Hellström, Emma Lazzeri, et al. «Recommendations for a Minimal Metadata Set to Aid Harmonised Discovery of Learning Resources», 9 June 2022. https://doi.org/10.15497/RDA00073.

'Integrating Research Infrastructures into Teaching: Recommendations and Best Practices'. n.d. Accessed 6 February 2025. https://doi.org/10.5281/zenodo.8114407.

Khan, A. F., Pedonese, G., Mallia, M., Frontini, F., Quochi, V., & Squadrito, E. (5 July 2024). Linguistic Linked Open Data for Humanists. Lisbon Summer School in Linguistics, Lisbon. Zenodo. https://doi.org/10.5281/zenodo.13897931.

Pedonese Giulia, Frontini Francesca, Ottaviani Roberta, Boschetti Federico, Spadi Alessia, Francalanci Lucia, Scognamiglio Alessia, Restaneo Pietro, Chaban Antonina, Striova Jana, Benassi Laura 'Materiali didattici come oggetti digitali FAIR: una metodologia condivisa per la formazione in H2IOSC' (AIUCD 2024), *in* a cura di: Di Silvestro, Antonio; Spampinato, Daria (2024) Me.Te. Digitali. Mediterraneo in rete tra testi e contesti, Proceedings del XIII Convegno Annuale

AIUCD2024. Catania: AIUCD, p. 595. ISBN 978-88-942535-8-0. DOI 10.6092/unibo/amsacta/7927. In: Quaderni di Umanistica Digitale.

'UNESCO Recommendation on Open Science - UNESCO Digital Library'. n.d. Accessed 1 December 2023. https://unesdoc.unesco.org/ark:/48223/pf0000379949.locale=en.

van der Lek, Iulianna, Darja Fišer, Francesca Frontini. «Introduction to Language Data: Standards and Repositories». In UPSKILLS Learning Content, 2023. (van der Lek et al., 2023 a) https://upskillsproject.eu/project/standards_repositories/.

van der Lek, Iulianna, Darja Fišer, Francesca Frontini, e Giulia Pedonese. «Introduzione ai Dati Linguistici: Standard e Archivi Digitali». 31 August 2024. https://doi.org/10.5281/zenodo.13911935.

van der Lek, Iulianna, Darja Fišer, Tanja Samardzic, Marko Simonovic, Stavros Assimakopoulos, Silvia Bernardini, Maja Milicevic Petrovic, e Genoveva Puskas. «Integrating Research Infrastructures into Teaching: Recommendations and Best Practices». Zenodo, 31 agosto 2023. (van der Lek et al., 2023 b) https://doi.org/10.5281/zenodo.8114407.

van der Lek, Iulianna, Francesca Frontini, Darja Fišer, Alexander König 2023. «Making the CLARIN Training Materials FAIR-By-Design», CLARIN Annual Conference 2023, Leuven 16 October - 18 October 2023. Deposited 20 January 2025. (van der Lek et al., 2023 c) https://doi.org/10.5281/zenodo.14699078.

Wilkinson, Mark, Michel Dumontier, IJsbrand Jan Aalbersberg, et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3, 160018 (2016). https://doi.org/10.1038/sdata.2016.18.