

Materiali didattici come oggetti digitali FAIR: una metodologia condivisa per la formazione in H2IOSC

Giulia Pedonese

Istituto di Linguistica
Computazionale “Antonio
Zampolli” (CNR-ILC), Italia
giulia.pedonese@cnr.it

Francesca Frontini

Istituto di Linguistica
Computazionale “Antonio
Zampolli” (CNR-ILC), Italia
francesca.frontini@cnr.
it

Roberta Ottaviani

Istituto di Linguistica
Computazionale “Antonio
Zampolli” (CNR-ILC), Italia
roberta.ottaviani@cnr.
it

Federico Boschetti

Istituto di Linguistica
Computazionale “Antonio
Zampolli” (CNR-ILC), Italia
federico.boschetti@cnr.i
t

Alessia Spadi

Istituto Opera del Vocabolario
Italiano (CNR-OVI), Italia
alessia.spadi@cnr.it

Lucia Francalanci

Istituto Opera del
Vocabolario Italiano (CNR-
OVI), Italia
lucia.francalanci@cnr.
it

Alessia Scognamiglio

Istituto per la Storia del Pensiero
Filosofico e Scientifico Moderno
(CNR-ISPF), Italia
alessia.scognamiglio@cnr
.it

Pietro Restaneo

Istituto per il Lessico
Intellettuale Europeo e Storia,
delle Idee (CNR-ILIESI), Italia
pietro.restaneo@cnr.it

Antonina Chaban

Istituto Nazionale di Ottica
(CNR-INO), Italia
antonina.chaban@cnr.it

Jana Striova

Istituto Nazionale di Ottica (CNR-INO), Italia
jana.striova@cnr.it

Laura Benassi

Istituto Nazionale di Ottica (CNR-INO), Italia
laura.benassi@cnr.it

Abstract

Il presente lavoro dettaglia la strategia per lo sviluppo di iniziative di formazione nell'ambito del progetto PNRR sviluppato dal CNR *Humanities and cultural Heritage Italian Open Science Cloud* (H2IOSC) e mira ad aprire alla comunità italiana di riferimento il processo di applicazione delle linee guida di design e di fruizione di moduli didattici che integrino l'uso delle Infrastrutture di Ricerca. In particolare, il contributo si sofferma sugli standard condivisi per la descrizione dei materiali didattici come oggetti digitali FAIR al fine di massimizzarne il riutilizzo in un'ottica *train the trainers* e sulla descrizione dei requisiti per l'implementazione dell'infrastruttura di *training*. Dopo aver descritto la strategia didattica (Sezione 2) e l'applicazione della metodologia FAIR-by-Design di Skills4EOSC ai materiali didattici preesistenti (Sezione 3), il lavoro descrive il processo di ideazione di due piattaforme con funzionalità coerenti ai requisiti degli oggetti didattici prendendo ad esempio il corso CLARIN *Introduction to Language Data: Standards and Repositories* tradotto e adattato in H2IOSC (Sezione 4).

Parole chiave: formazione, gestione dei dati, infrastrutture di ricerca, H2IOSC, principi FAIR

This paper details the strategy for the development of training initiatives in the framework of the PNRR project developed by the CNR *Humanities and cultural Heritage Italian Open Science Cloud* (H2IOSC) and aims to open up the process of applying the guidelines for the design and use of teaching modules integrating the use of Research Infrastructures to the Italian reference community. In particular, the contribution focuses on the shared standards for the description of teaching materials as FAIR digital objects to maximise their reuse in a train-the-trainer perspective and on the description of the requirements for the implementation of the training infrastructure. After describing the teaching strategy (Section 2) and the application of the Skills4EOSC FAIR-by-Design methodology to pre-existing learning materials (Section 3), the paper describes the process of designing two platforms with functionalities consistent with the requirements of learning objects, taking as an example the CLARIN course *Introduction to Language Data: Standards and Repositories* translated and adapted in H2IOSC (Section 4).

Keywords: Training, Data Management, Research Infrastructures, H2IOSC, FAIR Principles

1. La formazione nelle infrastrutture di ricerca

Il progetto *Humanities and cultural Heritage Italian Open Science Cloud* (H2IOSC) [1] intende creare un *cluster* federato dei servizi e delle risorse sviluppate dai nodi nazionali di quattro Infrastrutture di Ricerca (IR) per la Scienza Aperta¹ che fanno parte della *roadmap European Strategy Forum on Research Infrastructure* (ESFRI)² nel settore dell'innovazione sociale e culturale: *Digital Research Infrastructures for the Arts and Humanities* (DARIAH); *European Research Infrastructure for Heritage Science* (E-RIHS), *Common Language Resource and Technology Infrastructure* (CLARIN) e *Open Scholarly Communication in the European Research Area for Social Sciences and Humanities* (OPERAS).³ Il presente

¹ <https://unesdoc.unesco.org/ark:/48223/pf0000379949.locale=en>

² <https://roadmap2021.esfri.eu/landscape-analysis/section-1/social-cultural-innovation/>

³ H2IOSC rientra nell'ambito del PNRR Missione 4, "Istruzione e Ricerca" - Componente 2, "Dalla ricerca all'impresa" - Linea di investimento 3.1, "Fondo per la realizzazione di un sistema integrato di infrastrutture di ricerca e innovazione", finanziato dall'Unione europea – programma

contributo ha l'obiettivo di aprire la strategia di formazione sviluppata nell'ambito del progetto H2IOSC alla comunità scientifica in modo da favorire l'interscambio di obiettivi, metodologie e strumenti didattici volti a integrare l'uso delle IR all'interno dei programmi di istruzione superiore, universitaria e professionale.

Per il funzionamento delle IR intese come «strutture, risorse e servizi correlati che sono utilizzati dalla comunità scientifica per condurre ricerche di alto livello nei rispettivi campi del sapere» [2]: 197 la formazione è un aspetto cruciale in quanto costituisce un ponte fra le comunità scientifiche e ciò che le IR offrono, risultando strettamente connessa alle attività di disseminazione che promuovono l'uso di prodotti e servizi da parte di ricercatori, studenti e cittadini. Più in generale, le IR contribuiscono all'istruzione fornendo una formazione specializzata a studenti, ricercatori e professionisti del mondo accademico e industriale su metodi e tecnologie scientifiche all'avanguardia. Come riportato nel Piano Nazionale Infrastrutture di Ricerca (PNIR), le infrastrutture si configurano come luogo ideale per la formazione sul campo, sia accademica che industriale, con un conseguente impatto sociale positivo sulla formazione del capitale umano che potrà beneficiare degli *asset* fisici e tecnologici messi a disposizione dalle infrastrutture. Le IR possono inoltre contribuire a tirocini, tesi e stage universitari, oltre che farsi motore di formazione per nuove competenze necessarie alle infrastrutture stesse come vedremo nelle prossime sezioni.⁴ Per questo le iniziative di formazione hanno da sempre un posto di primo piano attraverso l'organizzazione di eventi come seminari, *workshop* e *summer school*, l'erogazione di borse di studio e di mobilità e la creazione di archivi digitali attraverso cui gli utenti possono accedere al materiale didattico, indicizzato secondo standard condivisi in quanto oggetto digitale in aderenza ai principi FAIR.⁵

A questo proposito si ricordano non solo le iniziative di infrastrutture disciplinari come CLARIN (con il suo *Learning Hub*)⁶ e DARIAH (con la piattaforma *DARIAH-Campus*),⁷ che forniscono l'accesso a risorse formative quali registrazioni di eventi, *slide deck* e tutorial indicizzati per tipologia di argomento o evento, ma anche le risorse dedicate al *training* da grandi progetti *cluster*

NextGenerationEU, di cui al Decreto Direttoriale MUR n. 0003264 del 28/12/2021 codice progetto IR0000029, CUP B63C22000730005. Sito ufficiale: <https://www.h2iosc.cnr.it/>

⁴ «Le IR sono strumento ideale per la formazione sul campo, sia che questa riguardi la formazione accademica che industriale, oltre a quella professionalizzante. Le IR sono fra le sedi ideali di tirocinio per lo svolgimento di dottorati di ricerca, sia tradizionali che industriali, per una serie di motivi: la loro intrinseca natura internazionale, la concentrazione di competenze e strumentazione unica, l'incontro con una comunità scientifica qualificata da anni di ricerca svolta in collaborazione. [...] Per questi stessi motivi, le IR possono anche essere sede ideale per tirocini di studenti dei corsi di Laurea, in particolare magistrale, che possono svolgere qui esperienze concrete molto importanti e formative, sia per l'acquisizione di conoscenze e competenze, sia per esperienza personale e acquisizione di competenze trasversali (soft skills)». *Decreto Ministeriale n.1082 del 10 settembre 2021- Piano Nazionale Infrastrutture di Ricerca (PNIR) 2021 – 2027.*

⁵ <https://www.go-fair.org/fair-principles/>

⁶ <https://www.clarin.eu/content/learning-hub>

⁷ I materiali di *training* sviluppati da DARIAH-EU sono disponibili su <https://campus.dariah.eu/>. All'interno del *framework*, le collezioni esterne sono cercabili tramite lo strumento DARIAH *Pathfinders* <https://campus.dariah.eu/resources/pathfinders>

europei fra i quali si segnala il *Training Discovery Toolkit*⁸ sviluppato nell'ambito del *Social Sciences and Humanities Open Cloud* (SSHOC).⁹ Inoltre, per potenziare le competenze trasversali degli utenti in termini di principi FAIR, Scienza Aperta e gestione dei dati della ricerca, il progetto *Skills for the European Open Science commons: creating a training ecosystem for Open and FAIR science* (Skills4EOSC),¹⁰ finanziato dal programma *Horizon Europe* della Commissione Europea e coordinato dal Consorzio GARR,¹¹ ricopre un ruolo d'avanguardia nello sviluppo di metodologie, attività e risorse condivise per unificare la formazione dei professionisti della ricerca che costituiranno il futuro bacino di utilizzatori dello *European Open Science Cloud* (EOSC).¹² Il progetto H2IOSC si inserisce pienamente in questo contesto dedicando l'intero *work package* 8 (WP8)¹³ alla formazione e basandosi sulle iniziative preesistenti per la Scienza Aperta e la gestione di dati FAIR per concentrarsi sull'ambito delle *Social Sciences and Humanities* (SSH) con casi d'uso specifici per le comunità afferenti.

2. La strategia di training H2IOSC

Il WP dedicato alla formazione è rivolto principalmente alle comunità di utenti italiani allo scopo di dotarli di capacità e competenze interdisciplinari specifiche del settore SSH, e formarli sulle risorse che le infrastrutture disciplinari possono offrire a livello nazionale e internazionale. Parte di questo obiettivo è quello di formare nuove figure professionali in grado di formare a loro volta le future generazioni su come integrare le IR nei metodi e nelle pratiche di ricerca delle rispettive discipline con una prospettiva *train the trainers*. Da questo deriva l'esigenza di disporre di materiali didattici modulari e facilmente riutilizzabili che i docenti possano integrare nei propri corsi di formazione a seconda delle necessità. Le unità operative coinvolte nel WP8 hanno sviluppato una strategia condivisa per il coordinamento delle attività di *training*, sia a livello di singole IR che di progetto, che comprende l'individuazione dei bisogni formativi delle diverse tipologie di utenti, l'implementazione di un'infrastruttura per il *training* che metta a disposizione una piattaforma per l'erogazione diretta di moduli formativi (sia in modalità sincrona che asincrona) e una per il deposito dei materiali didattici; e lo sviluppo di una metodologia comune per il design e l'adattamento di tali materiali.

Le tipologie di utenti oggetto della formazione sono state identificate in: ricercatori (esperti e all'inizio della loro carriera) con la necessità di integrare le proprie competenze per gestire correttamente i dati della ricerca; tecnologi che hanno bisogno di sviluppare dataset e strumenti da aprire alla comunità secondo i dettami della Scienza Aperta; docenti e formatori di ambito universitario e professionale che possano integrare l'uso delle IR nei loro curricula di insegnamento; studenti di corsi di laurea specialistica, magistrale e corsi di dottorato che devono

⁸ https://training-toolkit.sshopencloud.eu/entities?search=&f%5B0%5D=content_type%3Asource

⁹ <https://sshopencloud.eu/>

¹⁰ <https://www.skills4eosc.eu/>

¹¹ <https://www.garr.it/it/>

¹² <https://open-science-cloud.ec.europa.eu/>

¹³ Parlando di *Work Package* all'interno di un progetto complesso si intende un'unità di lavoro con obiettivi e finalità ben definiti all'interno del progetto stesso. Per maggiori dettagli sui *Work Package* del progetto H2IOSC si rimanda al sito del progetto <https://www.h2iosc.cnr.it/the-project/> [Ultimo accesso 2 Gennaio 2025].

essere esposti il prima possibile alle buone pratiche condivise; professionisti a supporto della ricerca come *data stewards* e *data curators*¹⁴ per cui è necessario trasmettere le proprie competenze a diversi tipi di utenti; e infine i professionisti delle istituzioni culturali (archivi, biblioteche, musei), che possono beneficiare dei servizi offerti dalle IR. Per sondare le necessità formative dei vari gruppi, il WP8, dedicato al *training*, lavora in stretto contatto con il WP dedicato all'analisi delle comunità di riferimento, che ha sviluppato un questionario e lo ha somministrato alla comunità scientifica [3]. A questo proposito, sono state implementate misure per il coordinamento con le varie associazioni, quali l'Associazione per l'Informatica Umanistica e la Cultura Digitale (AIUCD).

L'infrastruttura di *training* sviluppata dal progetto è composta da due piattaforme separate: *H2IOSC Training Environment*, cioè un *learning management system* [4] per l'erogazione di corsi in modalità sincrona e asincrona che permetterà ad H2IOSC e alle IR coinvolte di offrire formazione diretta sia al proprio personale che alle comunità di utenti; e la piattaforma *H2IOSC Training Library* allo scopo di conservare, condividere, citare e riutilizzare i materiali formativi come oggetti digitali FAIR dotati di licenza e di un sistema di gestione delle versioni, in modo da creare una comunità di formatori che condividano e riutilizzino i rispettivi materiali all'interno dei programmi accademici e di formazione. Questa infrastruttura dedicata alla formazione sarà una parte fondamentale del *Marketplace* H2IOSC [5] cioè la piattaforma che offrirà un punto di accesso unificato a risorse e servizi relativi alle Scienze Umane e Sociali a livello nazionale Italiano, progettata sul modello del *Social Sciences and Humanities Marketplace* [6]. Questi requisiti tecnici e funzionali sono stati tradotti in documenti di specifica utili all'implementazione di entrambe le piattaforme di *training* da parte del CNR - Istituto di Linguistica Computazionale "Antonio Zampolli" (CNR-ILC)¹⁵ con la collaborazione di ditte specializzate, per cui si rimanda alla Sezione 4 del presente articolo.

Quanto allo sviluppo di materiali di *training*, si è deciso di dedicare una prima fase all'adattamento di moduli didattici preesistenti, originati dalle iniziative di formazione già in atto nelle IR coinvolte, alle necessità degli utenti di H2IOSC. In un primo momento, l'erogazione di questi moduli è stata dedicata alla formazione del personale interno delle IR coinvolte in una prospettiva di potenziamento delle competenze. In un secondo momento sono stati promossi anche eventi rivolti a una comunità più ampia come, ad esempio, la partecipazione del nodo italiano di CLARIN (CLARIN-IT)¹⁶ alla *summer school* dal titolo *Digital Tools for Humanists*¹⁷ organizzata dall'Università di Pisa sotto la direzione del Laboratorio di Cultura Digitale e alla

¹⁴ La figura del *data steward* è una professionalità nascente di cui diverse iniziative a livello nazionale ed europeo hanno cercato di definire i contorni. Secondo la *Comunità Italiana dei Data Stewards* (CIDS), nata in seno al Centro di Competenza della *Italian Computing and Data Infrastructure*, il *data steward* è «un professionista dei dati della ricerca (...) che supporta i ricercatori nella gestione attenta e responsabile dei dati (e in generale di tutti i prodotti) della ricerca durante il loro ciclo di vita, ossia dalla pianificazione alla condivisione con il pubblico nell'ottica di realizzare una scienza sempre più aperta e trasparente». Fonte: https://open-science.it/it/article?rpk=313614&prs_sel=p_datasteward&tpc_sel=t_gestione_dei_dati_della_ricerca [Ultimo accesso 2 Gennaio 2025].

¹⁵ <https://www.ilc.cnr.it/>

¹⁶ <https://www.clarin-it.it/it>

¹⁷ <https://digitaltools.labcd.unipi.it/>

Lisbon Summer School in Linguistics,¹⁸ da cui è stato sviluppato un modulo didattico sui *Linguistic Linked Open Data* [7] attualmente in fase di traduzione e adattamento per il pubblico italiano. In parallelo, sono state messe in atto iniziative congiunte per consolidare le competenze trasversali del bacino di utenza H2IOSC e la formazione specifica per i membri del consorzio affinché possano a loro volta diventare formatori esperti nelle buone pratiche della Scienza Aperta e dei principi FAIR: a questo proposito, CLARIN-IT si è fatto carico di un'iniziativa per condividere l'uso della progettazione didattica retroattiva (*Backwards Instructional Design for Learning*) così come strutturata nel corso *FAIR-by-Design Methodology Training for Trainers*¹⁹ del progetto Skills4EOSC, che è alla base della stessa metodologia applicata da H2IOSC ai propri materiali didattici (Sezione 3). Tale corso è stato inizialmente diretto al personale interno della compagine H2IOSC e in seguito riproposto ad altri membri del consorzio. A coordinamento di tutte queste operazioni è stata sviluppata una metodologia di adattamento e design dei materiali didattici come oggetti digitali FAIR che verrà dettagliata nella prossima sezione.

3. Materiali didattici fair: una metodologia condivisa

Consapevoli che la produzione ed erogazione diretta di moduli didattici potrà soddisfare solo in parte le esigenze della comunità, H2IOSC vede nella costruzione di un ecosistema di FAIRificazione e condivisione di materiali didattici per il potenziamento e la trasmissione di competenze digitali una missione importante, anche al fine valorizzare le molte delle iniziative già intraprese da AIUCD, fra le quali citiamo il gruppo Scuola²⁰ e il ciclo di seminari di Aspettando AIUCD.²¹ Questa sezione illustra in particolare i principi alla base dello sviluppo della piattaforma di pubblicazione *H2IOSC Training Library*, dove andranno depositati dapprima tutti i materiali sviluppati dalle realtà coinvolte in H2IOSC e dai propri partner e, in seguito, anche i materiali della comunità scientifica più ampia, al fine di massimizzarne il riutilizzo. Secondo la strategia dettagliata nella sezione precedente, lo sviluppo dei moduli didattici ha preso le mosse da una prima fase di raccolta e sistematizzazione di materiali preesistenti (come ad esempio *slide deck*, infografiche, materiale audiovisivo, collezioni digitali ecc.). Ciò ha comportato l'ideazione di un flusso di lavoro capace sia di allineare i materiali con l'applicazione dei principi FAIR che di integrare i moduli già concepiti come FAIR, ma spesso descritti secondo standard differenti, in formati diversi e con licenze tra loro incompatibili. Per questo, nella definizione di una metodologia condivisa sono state delineate pratiche comuni per la descrizione dei materiali con metadati standard, l'applicazione di licenze adatte al riuso e l'attribuzione di identificativi univoci e persistenti (PID) che ne permettano la corretta citazione e l'attribuzione ai rispettivi autori.

Lo standard a cui si è fatto riferimento è la metodologia per lo sviluppo di materiali didattici FAIR-by-Design ideata dal progetto Skills4EOSC [8]. Questo *framework* garantisce che i materiali didattici siano allineati ai principi FAIR per aumentarne l'impatto e la riutilizzabilità all'interno della comunità scientifica e adotta un processo di progettazione didattica retroattiva dettagliato in sei fasi: preparazione, scoperta, design, sviluppo, pubblicazione e controllo della qualità. In accordo con queste raccomandazioni, l'unità minima modulare di riferimento è stata individuata

¹⁸ <https://www.ilc.cnr.it/events/lssl-2024/>

¹⁹ https://github.com/FAIR-by-Design-Methodology/FAIR-by-Design_ToT

²⁰ <https://www.aiucd.it/didattica-scuola/>

²¹ <https://www.aiucd.it/aspettando-aiucd2025/>

nel *learning object*, definito in Skills4EOSC come un pacchetto di una lezione, un'attività e una valutazione con un singolo obiettivo di apprendimento e un risultato di apprendimento concreto [8: 8]. In questo modo la singola lezione, corredata da materiali di accompagnamento e descritta con opportuni metadati, può essere utilizzata come base per la creazione di contenuti didattici più complessi. La metodologia di Skills4EOSC applica i principi FAIR ai *learning objects* adottando una doppia prospettiva: quella dello studente, che utilizza il prodotto finale su piattaforme di *e-learning*, e quella del formatore, che necessita di accedere a materiali didattici modificabili per adattarli alla propria attività di insegnamento. Per soddisfare entrambe le esigenze, il progetto H2IOSC ha sviluppato due tipologie di risorse digitali, entrambe strutturate come aggregati di singole lezioni intese come *learning objects*: i **corsi** ad accesso aperto in *e-learning* compatibili con i principali programmi di lettoscrittura, destinati agli studenti e resi disponibili su *H2IOSC Training Environment*; e **materiali didattici** modulari, scaricabili in formati aperti e facilmente editabili, a disposizione dei formatori su *H2IOSC Training Library*.²²

I metadati che descrivono ogni singolo *learning object* sono stati individuati nel *Minimal Metadata Set for Learning Resources* proposto dalla *Research Data Alliance* (RDA) [9] che stabilisce 14 campi suddivisi in 3 categorie di informazioni (*descriptive, access, educational*) ed è attualmente in fase di valutazione presso progetti come OpenPlato,²³ SSHOC Training Discovery Toolkit²⁴ e NI4OS Training Platform.²⁵ Si tratta di uno schema flessibile, aperto ad eventuali integrazioni a seconda delle necessità connesse, ad esempio, alle diverse prospettive di insegnamento (formali, professionali e informali) ed è volto a massimizzare la cercabilità dei dati senza appesantire il sistema dei descrittori, garantendo conformità e riusabilità del materiale esistente. Per la formazione H2IOSC, abbiamo adottato il modello già modificato da CLARIN [10], che ha investito nella formazione promuovendo pratiche standard e creando una comunità di formatori attraverso la *CLARIN Trainer's Network*. Tale modello ha modificato il set di metadati aggiungendo campi specifici, come *contributors, workload in ECTS, Persistent Identifier* (PID), *version date* e *standard citation* e lo ha quindi reso ancora più flessibile. In una prospettiva di applicazione dei principi FAIR dal punto di vista dei formatori, i metadati più importanti sono: quelli relativi a: 1) le parole chiave, che possono connettersi a vocabolari condivisi specifici per il settore disciplinare e quindi rendere il materiale facilmente indicizzabili e trovabili in un *repository* istituzionale; 2) le licenze utilizzate, che sono alla base delle possibilità di accesso e riutilizzo; 3) e infine i campi connessi a caratteristiche didattiche specifiche come *target group, learning outcome(s)* ed *expertise level*, che permettono di valutare la coerenza della lezione con le proprie finalità didattiche ed eventualmente pianificarne la rielaborazione.

In prospettiva del futuro deposito sulla *H2IOSC Training Library* si è deciso inoltre di optare per la conversione del materiale didattico preesistente nei formati più aperti disponibili, in modo da garantire la massima libertà e flessibilità di riutilizzo. Il formato non proprietario più versatile è sicuramente il *plaintext*, che è stato preferito laddove possibile, ad esempio rispetto a formati

²² Queste due diverse tipologie di oggetti sono state individuate anche con riferimento alle pratiche messe in atto da altri progetti europei, come ad esempio nel Deliverable 5.3: *Training Roadmap, Framework and Delivery of Training Library and Competence Centre* del progetto OSTrails, attualmente in fase di pubblicazione. Per una descrizione delle piattaforme, si veda la Sezione 4.

²³ <https://openplato.eu/blocks/catalog/list.php>

²⁴ <https://sshopencloud.eu/ssh-training-discovery-toolkit>

²⁵ <https://training.ni4os.eu/>

PDF (.ppt) o Moodle (.mbz). Per il design di nuovo materiale, che si tratti di creazione ex novo o di una rielaborazione di moduli preesistenti, inclusi quelli convertiti in *plaintext*, si prefigura l'utilizzo di *Markdown* come linguaggio *open-source* estensibile, intuitivo e facilmente convertibile in HTML, JSON, XML, YAML e molti altri formati. Lo stesso principio è alla base della scelta delle licenze,²⁶ che saranno più aperte possibile in considerazione delle regolamentazioni vigenti per la tipologia di dati in oggetto e della compatibilità delle licenze applicate al materiale riutilizzato. Un aspetto importante per facilitare il lavoro dei formatori e promuovere l'integrazione e il riutilizzo di *learning objects* è costituito dalle modalità di citazione dei materiali riutilizzati. A questo scopo, H2IOSC ha adottato pratiche formalizzate e condivise per permettere la corretta citazione e attribuzione dei propri materiali didattici, sia originali che derivati, considerati in tutto e per tutto come prodotti della ricerca. Fondamentale in questo senso è stato anche l'interscambio con le comunità scientifiche, quali la *Comunità Italiana dei Data Stewards* (CIDS)²⁷ nata in seno al Centro di Competenza della *Italian Computing and Data Infrastructure*²⁸ e in particolare con AIUCD, da sempre sensibile a queste tematiche, per la definizione di modalità di citazione, riuso e deposito chiare e di facile implementazione.

4. L'infrastruttura di training: le piattaforme H2IOSC

L'adattamento e l'applicazione della metodologia *FAIR-by-Design* di Skills4EOSC è stata un'operazione condotta in parallelo alla progettazione di un'infrastruttura dedicata alla formazione che potesse ospitare i corsi e i materiali didattici rendendoli fruibili per due diversi tipi di pubblico: coloro che intendono formarsi su competenze specifiche seguendo i corsi offerti dalle infrastrutture di H2IOSC (studenti, professionisti, ecc.) e i formatori che sono alla ricerca di materiale didattico riutilizzabile da adattare ai corsi di cui sono docenti. Ciò determina due diversi approcci nell'applicazione della metodologia, in quanto i principi FAIR di gestione dei dati devono essere adattati da un lato alle necessità degli studenti, che devono poter trovare, accedere, modificare materiale didattico scaricabile, editabile e fruibile su diversi supporti per garantire un'esperienza di apprendimento interattiva; e dall'altro devono garantire ai docenti la possibilità di depositare i propri prodotti e scoprire i materiali preesistenti in formato aperto e facilmente integrabile con più strumenti di progettazione didattica. Di seguito descriveremo le principali funzionalità delle piattaforme *H2IOSC Training Environment* e *H2IOSC Training Library* attraverso un esempio di applicazione della metodologia e le relative implicazioni a livello didattico. Mentre le specifiche tecniche dell'implementazione saranno dettagliate in un documento tecnico, il punto focale di questa sezione sarà il procedimento di FAIRificazione del corso *Introduction to Language Data: Standards and Repositories* [11] originariamente creato da CLARIN per il progetto UPSKILLS²⁹ e successivamente tradotto in italiano e riadattato da

²⁶ Preferibilmente Creative Commons: <https://creativecommons.it/chapterIT/index.php/license-your-work/>

²⁷ <https://open-science.it/-/data-steward-unite-breve-resoconto-del-kick-off-meeting-della-comunita-italiana-dei-data-steward>

²⁸ <https://www.icdi.it/it/attivita/competence-centre>

²⁹ Il progetto UPSKILLS (UPgrading the SKills of Linguistics and Language Students) è stato una partnership strategica Erasmus+ per l'istruzione superiore, mirata a identificare e colmare le lacune nelle competenze degli studenti di linguistica e discipline linguistiche attraverso lo sviluppo di nuovi componenti curriculari e materiali di supporto da integrare nei programmi di studio esistenti. Il progetto si è concluso il 31 agosto 2023. <https://upskillsproject.eu/>

H2IOSC con il titolo *Introduzione ai Dati Linguistici: Standard e Archivi Digitali* [12]. Questo caso specifico servirà a mostrare le implicazioni pratiche della metodologia al momento della pubblicazione per l'utilizzo da parte di due tipologie differenti di pubblico.

4.1. Pubblicare un corso per studenti

La piattaforma *H2IOSC Training Environment*³⁰ fornisce un ambiente interattivo per la fruizione di corsi in *e-learning* ed è stata progettata dal CNR-ILC. Dal lato della visualizzazione studente, la piattaforma permette la creazione di un *account* con un profilo personale che comprende i corsi seguiti e le certificazioni ottenute, la possibilità di iscriversi spontaneamente a corsi ad accesso libero o chiuso (in quest'ultimo caso con l'inserimento di un codice d'accesso fornito dal docente) e l'opportunità di interagire con gli altri utenti attraverso le funzionalità integrate di *chat* e *forum*. Lato docente invece, la piattaforma offre un sistema di *editing* veloce e intuitivo che permette di strutturare il materiale didattico secondo tre livelli gerarchici: il livello del **corso**, con una griglia che permette di comporre il *syllabus* generale seguendo lo schema di metadati RDA (Sezione 3); il livello del **modulo**, cioè un livello intermedio che permette di descrivere un'aggregazione coerente delle attività didattiche del corso secondo lo stesso modello di metadati; e infine il livello della **lezione**, che comprende i contenuti veri e propri (*slide deck*, video, infografiche, esercitazioni ecc.) e realizza un solo obiettivo didattico del corso. A questo livello, il docente può decidere di organizzare i contenuti in cartelle dedicate alle singole lezioni, di caricare semplicemente i documenti in formato testo, video e immagine, oppure di fornire il collegamento a una lezione online gestita da servizi di terzi (come ad esempio Big Blue Button, Teams, Zoom, Google Meets ecc.) che verrà integrato come un oggetto della piattaforma nell'apposita sezione "eventi" e nell'annesso calendario.

Nella prima metà del progetto, il consorzio CLARIN-IT, rappresentato dal suo membro fondatore e ospite del centro fornitori di servizi ILC4CLARIN³¹ presso il CNR-ILC, ha lavorato all'applicazione della metodologia Skills4EOSC al corso sviluppato da CLARIN per il progetto UPSKILLS dal titolo *Introduction to Language Data: Standards and Repositories*, già riutilizzabile secondo i principi FAIR e accessibile sulla piattaforma Moodle del progetto. Il corso, originariamente pensato per fornire a docenti e studenti di corsi di laurea in linguistica risorse e attività didattiche sull'uso di depositi di dati della ricerca nell'ambito del ciclo di vita dei dati linguistici, è stato tradotto in italiano e strutturato come un aggregato di file *Markdown* seguendo la metodologia Skills4EOSC. Per essere pubblicato come corso fruibile dagli studenti sulla piattaforma *H2IOSC Virtual Environment*, i materiali sono stati suddivisi in moduli e lezioni con una gerarchia che rispecchiasse il più possibile l'originaria struttura articolata in unità, lezioni e attività. Nella Figura 1 si può vedere l'esempio di una unità del corso originale, che comprende attività e presentazioni in HTML5 (H5P)³² sulla piattaforma Moodle del progetto UPSKILLS. Nelle figure seguenti la stessa unità, pubblicata come modulo didattico sulla *H2IOSC Training Environment*, è descritta da metadati standard (Figura 2) e comprende lezioni in formato testo interattivo nella visualizzazione degli studenti (Figura 3). In questo modo, l'esperienza di apprendimento è stata arricchita e resa maggiormente accessibile.

³⁰ Disponibile per l'iscrizione spontanea e gratuita all'indirizzo <https://h2iosc-training-platform.ilc4clarin.ilc.cnr.it/login>

³¹ <https://ilc4clarin.ilc.cnr.it/>

³² <https://h5p.org/>

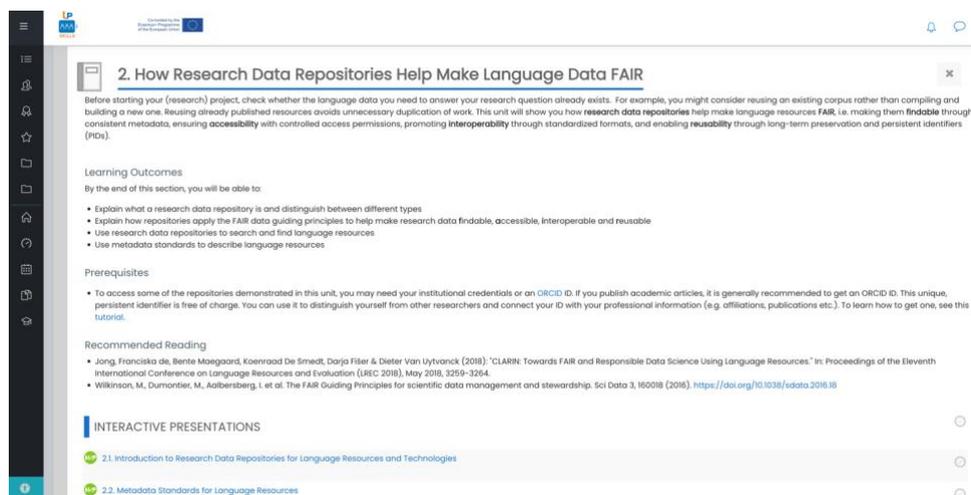


Figura 1 L'Unità 2 del corso Introduction to Language Data: Standards and Repositories sulla piattaforma Moodle del progetto UPSKILLS.

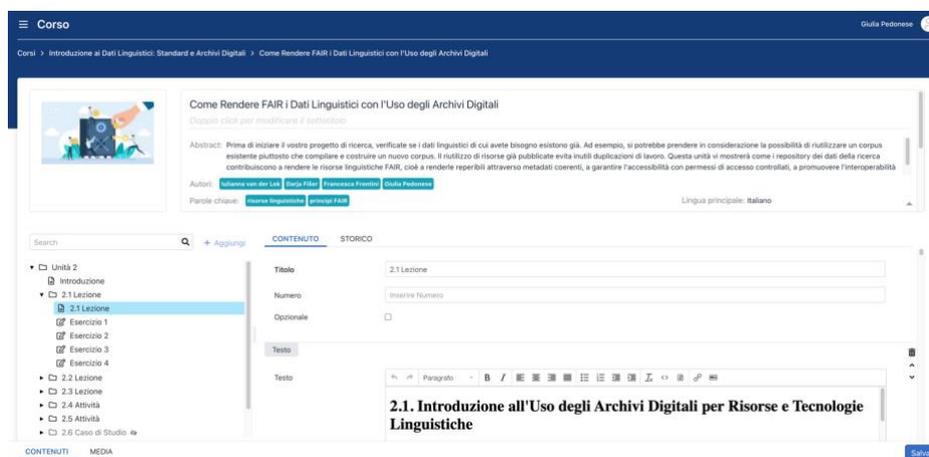


Figura 2 L'Unità 2 del corso Introduzione ai Dati Linguistici: Standard e Archivi Digitali in H2IOSC Training Environment: interfaccia di inserimento lato docente.

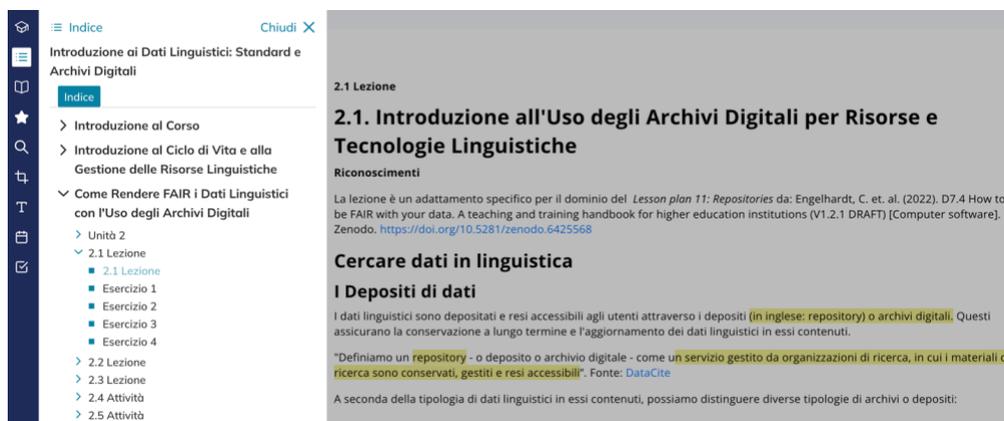


Figura 3 L'Unità 2 del corso Introduzione ai Dati Linguistici: Standard e Archivi Digitali in H2IOSC Training Environment: visualizzazione dello studente.

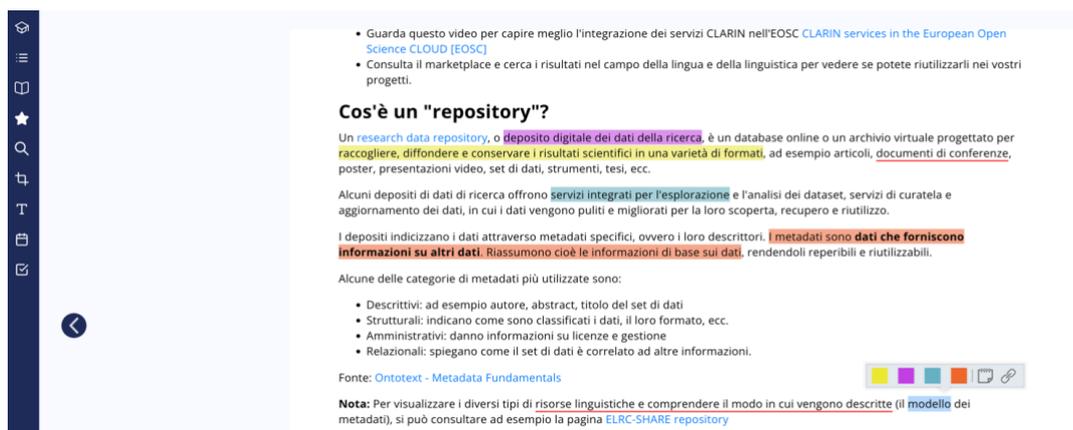
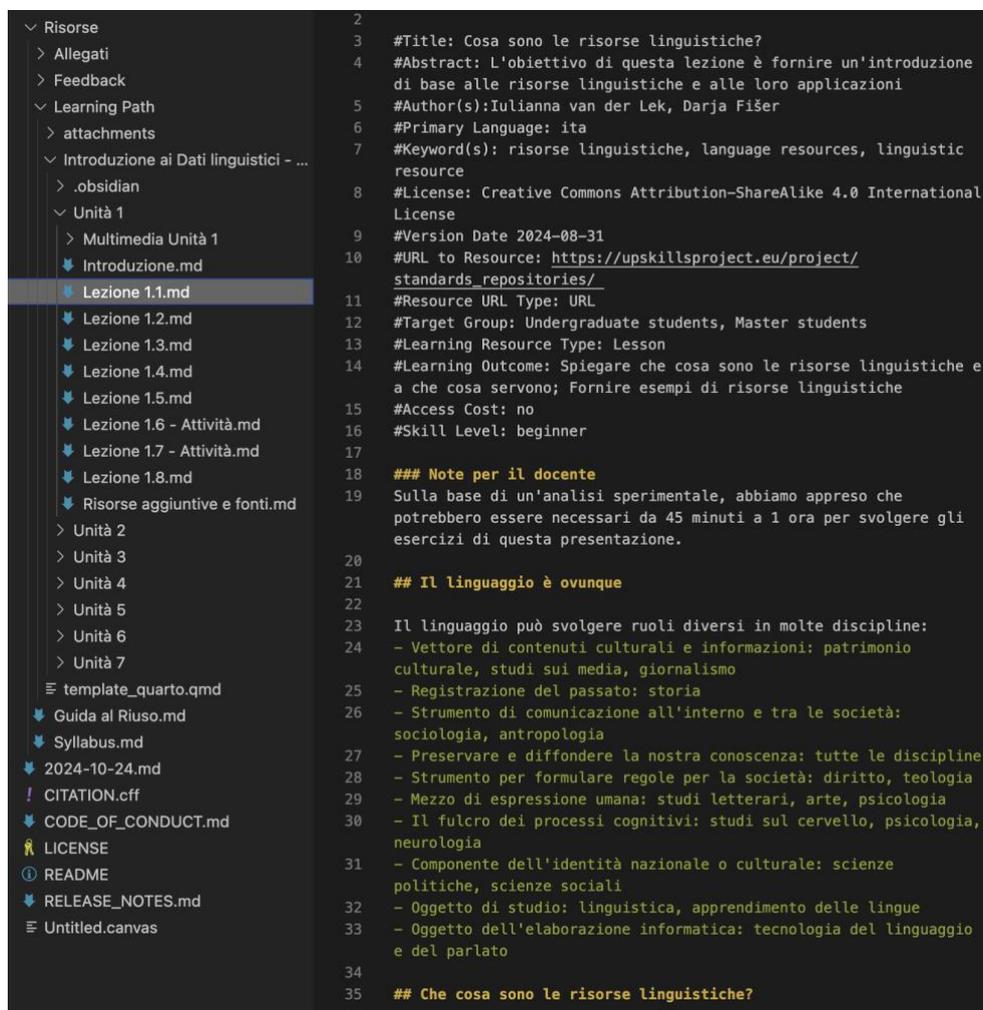


Figura 4 L'Unità 2 del corso Introduzione ai Dati Linguistici: Standard e Archivi Digitali in H2IOSC Training Environment: testo fluido con la possibilità di aggiungere sottolineature, collegamenti e note.

Mentre le necessità dello studente possono essere soddisfatte con la pubblicazione del materiale finale scaricabile, ad esempio, in formato testo o PDF, oppure come testo liquido che permette sottolineature e l'aggiunta di note o collegamenti (Figura 4), definire un aggregato di materiali didattici modulari accessibili e riutilizzabili da parte dei docenti presuppone un'elaborazione più approfondita che verrà descritta di seguito nella Sezione 4.2 con l'esempio dello stesso corso adattato da CLARIN.

4.2 Come rendere il materiale didattico riutilizzabile per i docenti

La H2IOSC Training Library è attualmente in corso di sviluppo da parte del CNR-ILC con il supporto di CENTRICA Srl. ed è una piattaforma che permette il deposito di materiale didattico modulare organizzato secondo la metodologia promossa da Skills4EOSC adattata alle necessità di H2IOSC (Sezione 3). Secondo le linee guida adottate, i materiali del corso *Introduzione ai Dati Linguistici: Standard e Archivi Digitali* sono stati convertiti in file .md (*Markdown*) e sono stati organizzati secondo una gerarchia che possa permettere all'utente di trovare e scaricare anche solo una parte dell'intero corso a seconda delle sue necessità. Per questo la piattaforma di deposito è stata ideata in modo da offrire un *backoffice* che faciliti le operazioni di deposito per chi vuole rendere disponibile il proprio materiale didattico. A questo scopo è stata predisposta un'interfaccia che permette di selezionare più contenitori, ordinati gerarchicamente dal più generale al più specifico: dal *learning path*, cioè la radice che si connette al catalogo di ognuna delle infrastrutture di ricerca coinvolte in H2IOSC, al *corso* e i suoi contenitori intermedi che permetteranno all'utente di selezionare solo una parte del materiale al momento del download: *sezione*, *unità* e *modulo*; fino all'ultimo contenitore che corrisponde all'unità didattica minima definita come *learning object* (Sezione 3). Gli unici contenitori obbligatori e che comprendono effettivamente degli oggetti sono: 1) il *learning path*, che ha la funzione di radice e contiene i documenti che accompagnano il materiale didattico facilitandone il riuso, come il file README, la citazione in formato .cff, il *code of conduct*, la *facilitation guide* e la licenza applicata al materiale, insieme al *syllabus* e a eventuali note di rilascio che accompagnano la gestione delle versioni successive; e 2) il *learning object*, cioè il nodo foglia che ha la funzione di contenere i materiali didattici in formato aperto e riproducibile. Ognuno di questi sei contenitori è descritto con il set di metadati RDA, con l'opzione di ereditare ed eventualmente modificare i campi di metadati già compilati nelle sezioni sovraordinate. Di seguito (Figura 5) si può vedere la struttura del materiale didattico del corso *Introduzione ai Dati Linguistici: Standard e Archivi Digitali* in un editor di testo *Markdown* (in questo caso Visual Studio Code) e un'anteprima di come questa struttura confluisce nelle funzionalità di caricamento della piattaforma (Figure 6 e 7).



```
2
3 #Title: Cosa sono le risorse linguistiche?
4 #Abstract: L'obiettivo di questa lezione è fornire un'introduzione
5 #Author(s):Iulianna van der Lek, Darja Fišer
6 #Primary Language: ita
7 #Keyword(s): risorse linguistiche, language resources, linguistic
8 #License: Creative Commons Attribution-ShareAlike 4.0 International
9 #Version Date 2024-08-31
10 #URL to Resource: https://upskillsproject.eu/project/standards\_repositories/
11 #Resource URL Type: URL
12 #Target Group: Undergraduate students, Master students
13 #Learning Resource Type: Lesson
14 #Learning Outcome: Spiegare che cosa sono le risorse linguistiche e
15 a che cosa servono; Fornire esempi di risorse linguistiche
16 #Access Cost: no
17 #Skill Level: beginner
18 ### Note per il docente
19 Sulla base di un'analisi sperimentale, abbiamo appreso che
20 potrebbero essere necessari da 45 minuti a 1 ora per svolgere gli
21 esercizi di questa presentazione.
22 ## Il linguaggio è ovunque
23 Il linguaggio può svolgere ruoli diversi in molte discipline:
24 - Vettore di contenuti culturali e informazioni: patrimonio
25 culturale, studi sui media, giornalismo
26 - Registrazione del passato: storia
27 - Strumento di comunicazione all'interno e tra le società:
28 sociologia, antropologia
29 - Preservare e diffondere la nostra conoscenza: tutte le discipline
30 - Strumento per formulare regole per la società: diritto, teologia
31 - Mezzo di espressione umana: studi letterari, arte, psicologia
32 - Il fulcro dei processi cognitivi: studi sul cervello, psicologia,
33 neurologia
34 - Componente dell'identità nazionale o culturale: scienze
35 politiche, scienze sociali
36 - Oggetto di studio: linguistica, apprendimento delle lingue
37 - Oggetto dell'elaborazione informatica: tecnologia del linguaggio
38 e del parlato
39 ## Che cosa sono le risorse linguistiche?
```

Figura 5 Esempio di struttura modulare del materiale didattico: Lezione 1.1 del corso Introduzione ai Dati Linguistici: Standard e Archivi Digitali.

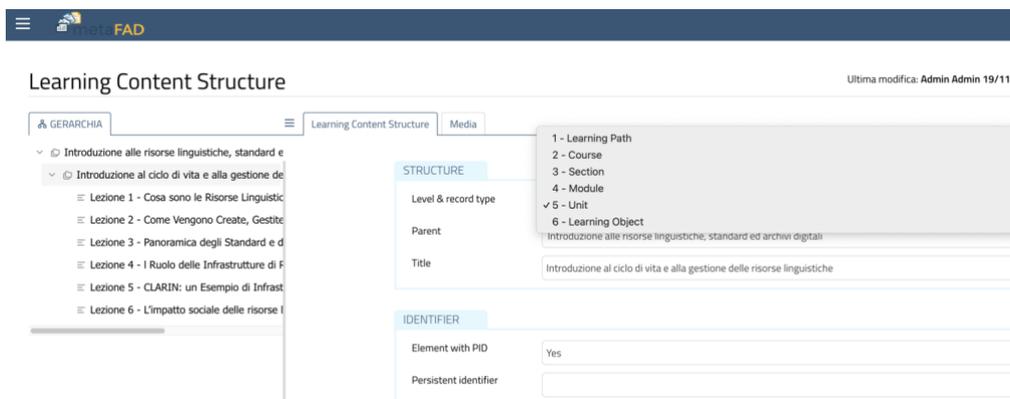


Figura 6 Esempio di caricamento dati nella H2IOSC Training Library: particolare della struttura dei contenuti didattici con selezione del livello contenitore (sezione Structure) e inserimento metadati, in particolare il PID (sezione Identifier).

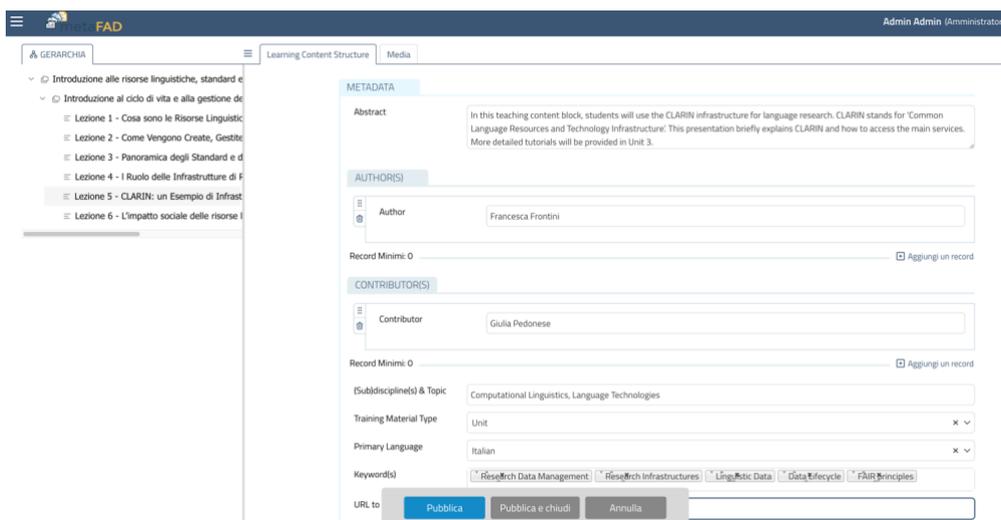


Figura 7 Esempio di caricamento dati nella H2IOSC Training Library: inserimento metadati secondo il set di RDA modificato da CLARIN.

La piattaforma inoltre supporta delle funzionalità di ricerca e scoperta del materiale didattico da parte di utenti interessati al riuso: tutti i materiali didattici saranno aperti in consultazione e saranno disponibili per il download previa autenticazione dell'utente attraverso un sistema di login federato. L'interfaccia di ricerca permetterà una ricerca semplice in linguaggio naturale, che ordinerà i risultati sulla base delle parole chiave inserite nell'apposita sezione *keyword* dei metadati. Un sistema a faccette permetterà di restringere ulteriormente la ricerca sulla base di una selezione più ampia di metadati, come la data di pubblicazione, la lingua, il nome dell'autore, il formato e l'argomento principale. Le figure 8-10 presentano rispettivamente un'anteprima dell'interfaccia

di ricerca, dei possibili risultati di una ricerca semplice, con i filtri a faccette in alto a destra che ne permettono il raffinamento e la *landing page* della risorsa selezionata, con il dettaglio dei contenuti e il pulsante per il *download*.

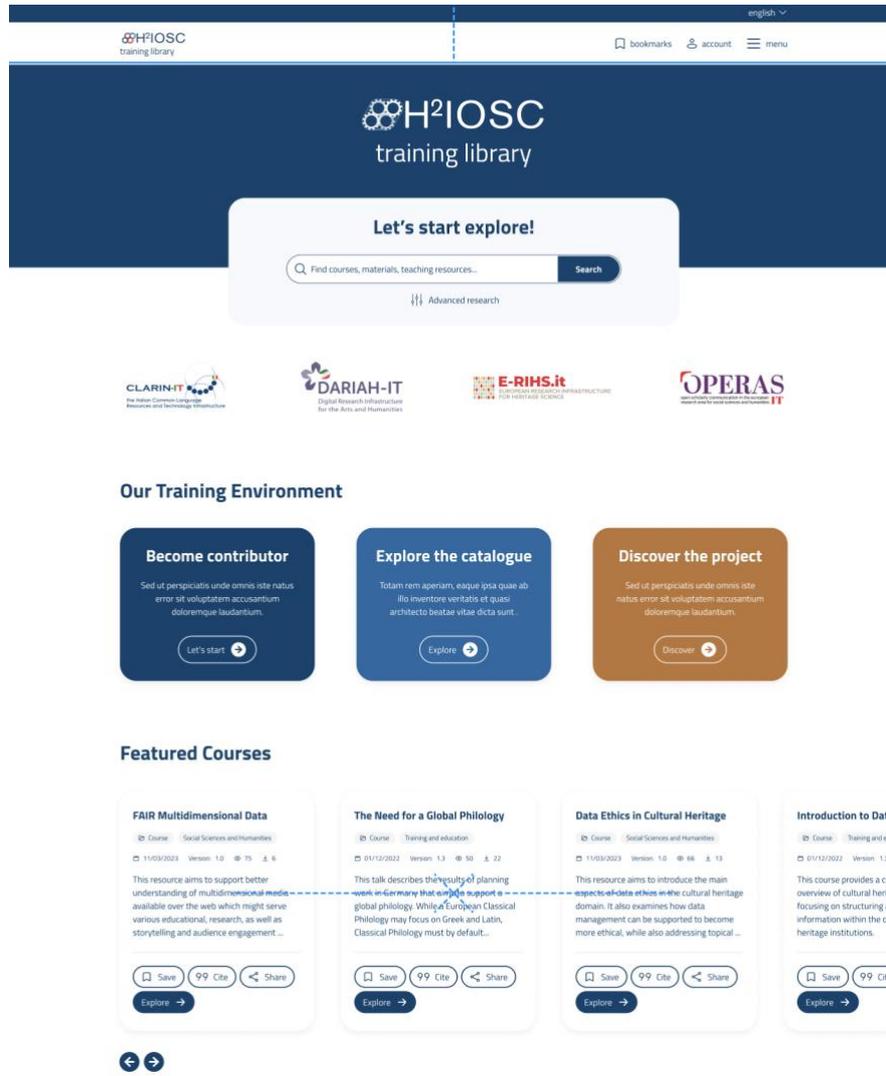


Figura 8 H2IOSC Training Library: esempio di interfaccia di ricerca.

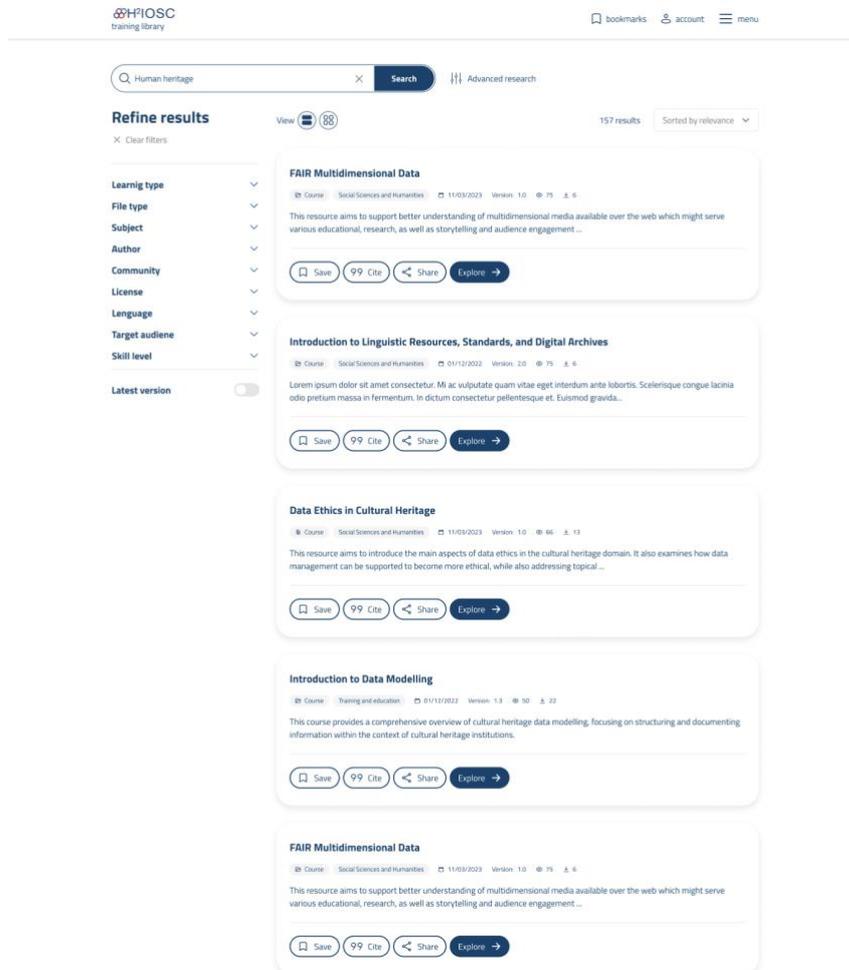


Figura 9 H2IOSC Training Library: esempio di interfaccia di ricerca con filtri a faccette.

The image shows a landing page for a course titled "Introduction to Linguistic Resources, Standards, and Digital Archives". The page has a dark blue header with the course title and a breadcrumb trail: Home / Social Sciences and Humanities / Introduction to Linguistic Resources, Standards, and Digital Archives. Below the header, there are course details: "Course: Social Sciences and Humanities", "11/03/2023", "Version: 2.0", and "6 lessons, 6 resources". A table lists course attributes: Company (Clarín), Target audience (Master's degree students), Skill level (Advanced), License (CC-BY-4.0), and Language (English). The page is divided into sections: "Details" (Authors: Lucrezia Grassi, Contributors: Giulio Romagnoli, Versions: Version 2.0, Topics: Computational Linguistics, Language Technologies, Persistent Identifier: https://publications.cnr.it/doi/10.5281/zenodo.75252, Keywords: Analyzing, Natural language processing, Academic publishing), "Abstract", "Learning outcome", "Prerequisite", and "In this Course". The "In this Course" section lists six lessons with "Explore" links. At the bottom, there are buttons for "Save", "Cite", "Share", and "Download all resources".

Figura 10 H2IOSC Training Library: esempio di landing page.

5. Risultati e prospettive future

L'adattamento della metodologia FAIR-by-design di Skills4EOSC e la sua applicazione ai materiali didattici sviluppati in H2IOSC ha permesso innanzitutto lo sviluppo coerente di due piattaforme di *training* orientate a rendere FAIR e perciò cercabili, accessibili, interoperabili e riutilizzabili i materiali didattici, costituendo una novità nel panorama delle infrastrutture di ricerca italiane ed europee.

I corsi già fruibili sulla piattaforma per studenti, *H2IOSC Training Environment*, comprendono i già citati corsi *Introduzione ai Dati Linguistici: Standard e Archivi Digitali* e *Linguistic Linked Open Data for Humanists* originato dalla partecipazione di CLARIN-IT alla *Lisbon Summer School in Linguistics* e un tutorial specifico su Weblicht [13], l'ambiente integrato per creare pipeline di trattamento automatico del linguaggio con strumenti resi disponibili dai vari centri CLARIN. Fra le prossime pubblicazioni, sono in programma: il corso *Introduzione alla Gestione dei Dati Orali* originato dalla formazione di CLARIN-IT per il progetto PRIN Corpus SIM³³ e una versione in italiano del corso organizzato da Skills4EOSC e CLARIN sull'adattamento della metodologia FAIR-by-design alla compagine di progetto, indirizzato ai formatori della rete H2IOSC [14]; il corso *La lessicografia web based all'Opera del Vocabolario Italiano: metodi e strumenti per la ricerca filologica e linguistica*, elaborato dal gruppo di lavoro di DARIAH-IT con tutorial sugli strumenti descritti, e i webinar³⁴ della serie didattica online *Current Topics in Heritage Science* che verranno adattati dal gruppo di lavoro di E-RIHS. Infine, vi saranno corsi dedicati all'uso dei servizi di H2IOSC come, ad esempio, il futuro *Marketplace*, e ad aspetti gestionali delle IR come l'*Infrastructure Management* per le *Digital Humanities*, con particolare focus sulla creazione di *data management plan* e aspetti legali ed etici connessi alla protezione dei dati digitali per le tipologie di dati rilevanti per H2IOSC.

6. Conclusioni

Le potenzialità di una metodologia condivisa per il design e l'adattamento di materiali didattici come prodotti della ricerca sono molteplici, primi fra tutti il superamento di divari metodologici e terminologici esistenti tra ambiti disciplinari diversi e l'agevolazione della formazione professionale nell'ambito di progetti nazionali e internazionali. Uno degli aspetti rilevanti della metodologia è la sua flessibilità nell'applicazione a diverse tipologie di pubblico: questo ha permesso ad esempio la progettazione di un modulo didattico rivolto ai docenti della scuola secondaria di secondo grado dal titolo *Introduzione alle Digital Humanities e all'Intelligenza Artificiale per la Scuola* organizzato dal gruppo AIUCD Scuola e l'associazione AICA,³⁵ ente accreditato per il rilascio di certificazioni ministeriali. Si tratta della terza edizione del corso nato su iniziativa di Dino Buzzetti (AIUCD). Quest'anno i materiali didattici verranno pubblicati anche sulla piattaforma H2IOSC Training Environment, oltre che sulla piattaforma di AICA.

Dato che stiamo attraversando un periodo storico di grande evoluzione tecnologica, si prevede che gli esperti di *Digital Humanities* svolgeranno un ruolo importante anche all'interno della formazione universitaria e professionale, con frequente necessità di aggiornamento. In un tale contesto, la creazione di un ecosistema per la condivisione di materiali didattici sarà un importante servizio alla comunità. In accordo con i principi della Scienza Aperta, il presente contributo apre alla comunità la metodologia sviluppata in seno al progetto H2IOSC per l'adattamento e il design di materiali didattici come *learning objects* modulari e riutilizzabili in modo

³³ Senectus Ipsa Morbus - *Spontaneous speech in healthy ageing*, progetto PRIN coordinato dalla prof.ssa Francesca Maria Dovetto <https://www.studiumanistici.unina.it/ricerca/progetti-di-ricerca/progetti-nazionali/progetti-prin/>

³⁴ Si tratta di materiale originato da incontri mensili di formazione online su aspetti fondamentali dell'*Heritage Science*: <https://www.iperionhs.eu/lectures-series/>

³⁵ <https://www.aicanet.it/>

da massimizzare l'impatto delle iniziative di formazione che integrano l'uso delle Infrastrutture di Ricerca.

Ringraziamenti

Gli autori e le autrici dichiarano di avere utilizzato l'intelligenza artificiale generativa, in particolare ChatGPT, allo scopo di riassumere le informazioni, tradurre e migliorare il testo del presente articolo.

Progetto H2IOSC - Humanities and cultural Heritage Italian Open Science Cloud finanziato dall'Unione europea NextGenerationEU - Piano Nazionale di Ripresa e Resilienza (PNRR) - Missione 4 "Istruzione e Ricerca" Componente 2 "Dalla ricerca all'impresa" Linea di Investimento 3.1 "Fondo per la realizzazione di un sistema integrato di infrastrutture di ricerca e innovazione" Azione 3.1.1 "Creazione di nuove IR o potenziamento di quelle esistenti che concorrono agli obiettivi di Eccellenza Scientifica di Horizon Europe e costituzione di reti" - Codice progetto IR0000029 - CUP B63C22000730005. Soggetto attuatore CNR.

Riferimenti

- [1] Degl'Innocenti, Emiliano, Monica Monachini, Alberto Bucciero, Enrico Pasini, Bruno Fanini, e Francesca Frontini. «H2IOSC: Humanities and Heritage Open Science Cloud». In *La memoria digitale: forme del testo e organizzazione della conoscenza. Atti del XII Convegno Annuale AIUCD*, a cura di Emmanuela Carbé, Gabriele Lo Piccolo, Alessia Valenti, e Francesco Stella, 63–64, 2023. <https://iris.unive.it/retrieve/0f226d38-e332-418b-9b14-d5558d1a0d9d/AIUCD2023.pdf>.
- [2] Frontini, Francesca, e Monica Monachini. «Infrastrutture digitali per le scienze umane e sociali». In *Digital Humanities. Metodi, strumenti, saperi*, 197–213. Roma: Carocci, 2023. <https://www.carocci.it/prodotto/digital-humanities>.
- [3] Luzietti, Roberta Bianca, et al. «Digital Humanities and Heritage Science: moving from landscaping to a dynamic research observatory in an Open Science Cloud». In *Proceedings of the AIUCD 2024*, pp. 559-555. <https://www.iris.unict.it/retrieve/69fe4e9f-bdd2-4a15-b092-ee87002baa25/2024%20AIUCD%20Paves-e.pdf>.
- [4] Turnbull, Darren, Ritesh Chugh, e Jo Luck. «Learning Management Systems, An Overview». In *Encyclopedia of Education and Information Technologies*, a cura di Arthur Tatnall, 1052–58. Cham: Springer International Publishing, 2020. https://doi.org/10.1007/978-3-030-10576-1_248.
- [5] Sichera, Pietro, Cristina Marras, e Enrico Pasini. «ORCHESTRAZIONE API PER WORKFLOW APPLICATIVI NELL'AMBITO DELLE DIGITAL HUMANITIES». Zenodo, 19 novembre 2024. <https://doi.org/10.5281/zenodo.14187534>.
- [6] Barbot, Laure, Edward J. Gray, Frank Fischer, Matej Durco, Alexander König, Marie Puren, Stefan Buddenbohm, Cesare Concordia, e Klaus Illmayer. «The SSH Open Marketplace: A Multi-Voiced Story». Report. Zenodo, 2022. <https://doi.org/10.5281/zenodo.6580302>.
- [7] Khan, Anas Fahad, Giulia Pedonese, Michele Mallia, Frontini Frontini, Valeria Quochi, e Elisa Squadrito. «Linguistic Linked Open Data for Humanists». 5 luglio 2024. <https://doi.org/10.5281/zenodo.13897931>.

- [8] Filiposka, Sonja, Dominique Green, Anastas Mishev, Vojdan Kjorveziroski, Andrea Corleto, Eleonora Napolitano, Gabriella Paolini, et al. «D2.2 Methodology for FAIR-by-Design Training Materials», 31 agosto 2023. <https://zenodo.org/records/8305540>.
- [9] 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 giugno 2022. <https://doi.org/10.15497/RDA00073>.
- [10] Lek, Iulianna van der, Francesca Frontini, Darja Fiser, e Alexander König. «Making the CLARIN Training Materials FAIR-By-Design». 20 gennaio 2025. <https://doi.org/10.5281/zenodo.14699078>.
- [11] Lek, Iulianna van der, Darja Fišer, Francesca Frontini. «Introduction to Language Data: Standards and Repositories». In UPSKILLS Learning Content, 2023. https://upskillsproject.eu/project/standards_repositories/.
- [12] Lek, Iulianna van der, Darja Fišer, Francesca Frontini, e Giulia Pedonese. «Introduzione ai Dati Linguistici: Standard e Archivi Digitali». 31 agosto 2024. <https://doi.org/10.5281/zenodo.13911935>.
- [13] Pedonese, Giulia. «Le funzionalità di Weblicht: ricerca e annotazione linguistica di un testo attraverso i servizi di CLARIN». 26 giugno 2024. <https://doi.org/10.5281/zenodo.14236201>.
- [14] Filiposka, Sonja, Anastas Mishev, Athina Anastasopoulou, Francesca Frontini, Giulia Pedonese, e Iulianna van der Lek. «Fair-by-design Training for the CLARIN Community». 1 ottobre 2024. <https://doi.org/10.5281/zenodo.13867278>.