

# Dai Materiali Didattici alle Piattaforme FAIR: Costruire un'Infrastruttura di Training in H2IOSC

Giulia Pedonese<sup>1</sup>, Francesca Frontini<sup>2</sup>, Roberta Ottaviani<sup>3</sup>, Federico Boschetti<sup>4</sup>, Alessia Spadi<sup>5</sup>, Lucia Francalanci<sup>6</sup>, Alessia Scognamiglio<sup>7</sup>, Pietro Restaneo<sup>8</sup>, Antonina Chaban<sup>9</sup>, Jana Striova<sup>10</sup>, Laura Benassi<sup>11</sup>

<sup>1</sup> Istituto di Linguistica Computazionale "Antonio Zampolli" (CNR-ILC), Italia – [giulia.pedonese@cnr.it](mailto:giulia.pedonese@cnr.it)

<sup>2</sup> Istituto di Linguistica Computazionale "Antonio Zampolli" (CNR-ILC), Italia – [francesca.frontini@cnr.it](mailto:francesca.frontini@cnr.it)

<sup>3</sup> Istituto di Linguistica Computazionale "Antonio Zampolli" (CNR-ILC), Italia – [roberta.ottaviani@cnr.it](mailto:roberta.ottaviani@cnr.it)

<sup>4</sup> Istituto di Linguistica Computazionale "Antonio Zampolli" (CNR-ILC), Italia – [federico.boschetti@cnr.it](mailto:federico.boschetti@cnr.it)

<sup>5</sup> Istituto Opera del Vocabolario Italiano (CNR-OVI), Italia – [alessia.spadi@cnr.it](mailto:alessia.spadi@cnr.it)

<sup>6</sup> Istituto Opera del Vocabolario Italiano (CNR-OVI), Italia – [lucia.francalanci@cnr.it](mailto:lucia.francalanci@cnr.it)

<sup>7</sup> Istituto per la Storia del Pensiero Filosofico e Scientifico Moderno (CNR-ISPF), Italia – [alessia.scognamiglio@cnr.it](mailto:alessia.scognamiglio@cnr.it)

<sup>8</sup> Istituto per il Lessico Intellettuale Europeo e Storia, delle Idee (CNR-ILIESI), Italia – [pietro.restaneo@cnr.it](mailto:pietro.restaneo@cnr.it)

<sup>9</sup> Istituto Nazionale di Ottica (CNR-INO), Italia – [antonina.chaban@cnr.it](mailto:antonina.chaban@cnr.it)

<sup>10</sup> Istituto Nazionale di Ottica (CNR-INO), Italia – [jana.striova@cnr.it](mailto:jana.striova@cnr.it)

<sup>11</sup> Istituto Nazionale di Ottica (CNR-INO), Italia – [laura.benassi@cnr.it](mailto:laura.benassi@cnr.it)

## ABSTRACT (ITALIANO)

Questo contributo si propone di illustrare la progettazione e lo sviluppo di un'infrastruttura di formazione innovativa per le Scienze Umane e Sociali, basata sui principi FAIR e sulla promozione della Scienza Aperta, nell'ambito del progetto Humanities and cultural Heritage Italian Open Science Cloud (H2IOSC). L'obiettivo principale è la creazione di un ecosistema integrato che renda i materiali didattici facilmente reperibili, accessibili, interoperabili e riutilizzabili. A tal fine, sono state implementate due piattaforme: *H2IOSC Virtual Environment*, dedicata all'erogazione di corsi e risorse per studenti, e *H2IOSC Training Library*, un deposito per la conservazione e la condivisione di materiali didattici modulari. Entrambe le piattaforme si basano sulla metodologia "FAIR-by-Design" raccomandata dal progetto Skills4EOSC, che struttura il processo educativo in sei fasi, garantendo standard elevati di metadato e l'uso di formati aperti. Con l'implementazione di queste piattaforme, i cui servizi saranno resi disponibili a un livello di aggregazione più alto nel *Marketplace* di H2IOSC, il progetto intende favorire un approccio scalabile e sostenibile alla formazione, promuovendo al contempo la collaborazione tra docenti e studenti.

**Parole chiave:** formazione; gestione dei dati; infrastrutture di ricerca; principi FAIR; Scienza Aperta.

## ABSTRACT (ENGLISH)

*From Learning Materials to FAIR Platforms: Building a Training Infrastructure in H2IOSC.*

This paper aims to illustrate the design and development of an innovative training infrastructure for the Social Sciences and Humanities based on FAIR principles and the promotion of Open Science, within the Humanities and cultural Heritage Italian Open Science Cloud (H2IOSC) project. The main objective is the creation of an integrated ecosystem that makes learning materials easily findable, accessible, interoperable and reusable. To this end, two platforms have been implemented: *H2IOSC Virtual Environment*, dedicated to the delivery of courses and student resources, and *H2IOSC Training Library*, a repository for the storage and sharing of modular teaching materials. Both platforms are based on the 'FAIR-by-Design' methodology recommended by the Skills4EOSC project, which structures the educational process in six steps, ensuring high standards of metadata and the use of open formats. By implementing these platforms, whose services will be made available at a higher aggregation level in the H2IOSC Marketplace, the project intends to foster a scalable and sustainable approach to education, while promoting collaboration between teachers and students.

**Keywords:** training; data management; research infrastructures; FAIR principles; Open Science.

## 1. CONTESTO E OBIETTIVI

Il contributo si inserisce nell'ambito della formazione gestita dal *Work Package 8* (WP8) del progetto *Humanities and cultural Heritage Italian Open Science Cloud*<sup>1</sup> (H2IOSC), progetto PNRR che mira a creare un cluster federato di servizi e risorse sviluppati dai nodi nazionali di quattro infrastrutture di ricerca europee per la Scienza Aperta nel settore delle Scienze Umane e Sociali: CLARIN, DARIAH, E-RHIS e OPERAS (Degl'Innocenti, 2023). Nell'ambito della formazione, le iniziative sono spesso disperse su piattaforme differenti, non sempre descritte con metadati specifici e accessibili al pubblico. Per questo il

<sup>1</sup> <https://www.h2iosc.cnr.it/>

progetto H2IOSC ha previsto l'implementazione di una metodologia condivisa (Pedonese, 2024) per strutturare i materiali didattici in modo che siano utilizzabili in accordo con i principi FAIR di gestione dei dati e di un'infrastruttura composta da due piattaforme: *H2IOSC Virtual Environment* per l'erogazione e la fruizione del training da parte degli utenti (identificati come studenti universitari, ricercatori, docenti delle scuole secondarie, professionisti del settore culturale e professionisti a sostegno della ricerca come Data Stewards e Data librarians); e *H2IOSC Training Library*, un *repository* specifico per la pubblicazione di materiali didattici modulari in formato editabile che permette l'assegnazione di *Persistent identifiers* (PID), licenze standard (ad esempio Creative Commons) e una gestione integrata per l'aggiornamento delle versioni, in modo che materiali didattici siano accessibili e riutilizzabili per docenti universitari e formatori professionali. Il contributo dettaglia la progettazione delle due piattaforme e il loro design integrato secondo la metodologia condivisa.

## 2. LA METODOLOGIA DIDATTICA CONDIVISA DI H2IOSC

Il progetto H2IOSC si è allineato al progetto Skills4EOSC<sup>2</sup> adottando la metodologia "FAIR-by-Design" (Filiposka, 2023) per lo sviluppo di risorse didattiche. Questa metodologia, basata sui principi FAIR (*Findable, Accessible, Interoperable, Reusable*), segue un processo in sei fasi (preparazione, scoperta, design, sviluppo, pubblicazione, e controllo qualità) e integra strumenti pratici come checklist e modelli per garantire un approccio strutturato e scalabile. In particolare, H2IOSC ha adottato la definizione di "*learning object*" proposta da Skills4EOSC, considerandolo come una lezione modulare caratterizzata da obiettivi di apprendimento specifici che include non solo attività, ma anche strumenti di valutazione delle competenze acquisite. Inoltre, tutte le risorse didattiche di H2IOSC sono descritte con il *Minimal Metadata Set for Learning Resources* proposto dalla Research Data Alliance (Hoebelheinrich, 2022) così come arricchito dal modello CLARIN (van der Lek, 2025), che aggiunge campi come ad esempio *contributors*, *workload in ECTS* e *Persistent Identifier* (PID) per migliorare l'utilizzabilità e la documentazione delle risorse. Per garantire la massima cercabilità dei materiali didattici, si è deciso di compilare il set di metadati di ogni risorsa inizialmente in lingua inglese (per mantenere l'allineamento con Skills4EOSC), prevedendo comunque una gestione multilingue dei metadati, utilizzando anche le risorse terminologiche multilingui di CLARIN (Frontini, 2021; Borek, 2020). Per una maggiore riusabilità, H2IOSC ha scelto di convertire il materiale esistente in formati aperti come *plaintext*, per garantire flessibilità e riuso, e di utilizzare Markdown per la creazione di nuovi contenuti. Le licenze saranno il più aperte possibile, in linea con le raccomandazioni di Skills4EOSC, per favorire il riutilizzo. Inoltre, H2IOSC ha sviluppato pratiche di citazione formalizzate per consentire la tracciabilità e il riuso dei materiali, lavorando in stretta collaborazione con comunità scientifiche come AIUCD per promuovere standard condivisi e sostenibili. Uno dei punti cardine della metodologia di Skills4EOSC è l'applicazione dei principi FAIR alla progettazione didattica sotto due punti di vista: da un lato quello dello studente, che fruisce del prodotto finito su piattaforme di e-learning; dall'altro quello del formatore, che ha la necessità di poter trovare, accedere e riutilizzare materiali didattici in formati editabili a seconda delle proprie esigenze. Per questo il progetto H2IOSC ha previsto la produzione di due diverse tipologie di oggetti digitali: corsi ad accesso aperto in modalità e-learning rivolti agli studenti in formati compatibili con i maggiori programmi di lettoscrittura e materiali didattici modulari disponibili per il download in formati aperti e modificabili secondo le necessità dei formatori. Per la fruizione di questi due tipi di oggetti digitali, il progetto H2IOSC ha previsto l'implementazione di due diverse piattaforme: un *learning management system* chiamato *H2IOSC Training Environment* per l'accesso ai corsi da parte degli studenti e un *repository* di materiali didattici per formatori chiamato *H2IOSC Training Library*.

## 3. H2IOSC TRAINING ENVIRONMENT

La piattaforma *H2IOSC Training Environment* è progettata per la pubblicazione e la fruizione di materiale didattico strutturato in corsi, moduli e lezioni. I corsi sono destinati alle persone (studenti, professionisti, ricercatori ecc.) che intendono acquisire o migliorare le proprie competenze nell'ambito delle scienze umane e sociali, con particolare attenzione agli strumenti messi a disposizione dalle infrastrutture di ricerca afferenti ad H2IOSC: per questo i nodi italiani di CLARIN, DARIAH, E-RIHS e OPERAS hanno prodotto un'offerta di corsi in autoformazione relativi, ad esempio, alla gestione di dati linguistici secondo i

---

<sup>2</sup> <https://www.skills4eosc.eu/>

principi FAIR e Open Science, alle Digital Humanities per la didattica e all'uso di modelli per comprendere la pittura a olio e il patrimonio artistico.<sup>3</sup>

La piattaforma è stata sviluppata da ETT S.p.A. in collaborazione con l'Istituto CNR di Linguistica Computazionale "Antonio Zampolli" di Pisa (CNR-ILC) sulla base di una progettazione condivisa dalle quattro infrastrutture in modo da offrire un'esperienza completa sia per gli studenti sia per i docenti afferenti ai diversi domini di ricerca.

Dal lato *front-end*, la piattaforma consente agli studenti la creazione di un profilo personale che raccoglie i corsi seguiti e le certificazioni conseguite, l'iscrizione spontanea a corsi, sia ad accesso libero sia ad accesso chiuso (in quest'ultimo caso mediante un codice fornito dal docente) e l'interazione con altri utenti tramite funzionalità integrate di chat e forum.

Dal *back-end*, la piattaforma mette a disposizione dei docenti un sistema di *editing* intuitivo per organizzare il materiale didattico su tre livelli gerarchici: 1) il livello corso, che consente di definire il *syllabus* generale utilizzando una griglia conforme al modello di metadati RDA (Sezione 2); 2) il livello modulo, che rappresenta un livello intermedio dedicato alla descrizione di un sottogruppo di attività didattiche seguendo il medesimo schema di metadati del corso; 3) il livello dei materiali didattici, che riguarda i contenuti specifici di ciascun modulo, da organizzare in cartelle per le singole lezioni oppure da caricare direttamente come file in formato testo, video o immagine; in alternativa, è possibile integrare collegamenti a lezioni online tramite servizi esterni (es. Big Blue Button, Teams, Zoom, Google Meets), che verranno gestiti come oggetti della piattaforma nella sezione "eventi" e nel calendario integrato. Grazie a queste funzionalità, *H2IOSC Training Environment* rappresenta una soluzione modulare che rende il processo di insegnamento-apprendimento interattivo e versatile.

#### 4. H2IOSC LIBRARY

*H2IOSC Training Library* è una piattaforma in fase di sviluppo grazie alla collaborazione tra M.E.T.A. S.r.l. e CNR-ILC. Si tratta di un repository, attualmente a nostra conoscenza unico nel suo genere, esclusivamente dedicato al deposito FAIR di materiale didattico modulare, organizzato secondo la metodologia proposta da Skills4EOSC opportunamente adattata alle esigenze di H2IOSC (Sezione 2). Il design della piattaforma è stato testato su un materiale didattico di esempio, cioè il corso *Introduction to Language Data: Standards and Repositories* (van der Lek, 2023) originariamente sviluppato da CLARIN per il progetto UPSKILLS<sup>4</sup>, in seguito tradotto e riadattato per la comunità italiana nell'ambito di H2IOSC. Seguendo le linee guida adottate, i contenuti del corso tradotto, dal titolo *Introduzione ai Dati Linguistici: Standard e Archivi Digitali* (van der Lek, 2024), sono stati convertiti in file .md (*Markdown*) e strutturati in una gerarchia che consente agli utenti di individuare e scaricare singole parti del corso in base alle proprie necessità. La piattaforma, infatti, è stata progettata per offrire un *backoffice* che semplifica il deposito del materiale didattico, supportando chi desidera rendere i propri contenuti disponibili.

L'interfaccia della *H2IOSC Training Library* consente di selezionare vari contenitori organizzati gerarchicamente, dal più generale al più specifico: il *learning path*, cioè la radice che si collega al catalogo delle infrastrutture di ricerca partecipanti a H2IOSC; il *Corso*, che è il contenitore più ampio, e relativi contenitori intermedi: *sezione*, *unità* e *modulo*, che permettono di selezionare porzioni specifiche di materiale da scaricare, fino al contenitore minimo che rappresenta la porzione didattica più elementare, chiamata *learning object* (Sezione 2). Gli unici oggetti obbligatori per il deposito sono: 1) il *learning path*, che funge da radice e include documenti essenziali per facilitare il riuso del materiale didattico (quali 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 i metadati già compilati nelle sezioni sovraordinate.

La piattaforma include inoltre un'interfaccia di ricerca con funzionalità avanzate per la scoperta del materiale didattico depositato che si rivolge a utenti interessati al riuso. Tutti i materiali didattici saranno consultabili liberamente, mentre il *download* sarà possibile solo previa autenticazione tramite un sistema di autenticazione federata. L'interfaccia di ricerca offrirà la possibilità di effettuare ricerche semplici in linguaggio naturale, con i risultati ordinati in base alle parole chiave inserite nella sezione apposita dei metadati. Inoltre, un sistema a faccette permetterà di affinare ulteriormente i risultati della ricerca,

<sup>3</sup> Il catalogo completo e aggiornato dell'offerta formativa disponibile sulla H2IOSC Training Environment è disponibile all'indirizzo: <https://www.h2iosc.cnr.it/training-infrastructure/>

<sup>4</sup> <https://upskillsproject.eu/>

utilizzando una gamma più ampia di criteri come la data di pubblicazione, la lingua, il nome dell'autore, il formato e l'argomento principale.

Le funzionalità di entrambe le piattaforme, relative alla fruizione, al deposito e alla scoperta di materiale didattico modulare puntano a massimizzarne il riutilizzo in accordo con i principi FAIR di gestione dei dati e le buone pratiche di Scienza Aperta. Nella prossima sezione si anticiperanno le modalità di accesso integrato a queste funzionalità attraverso il Marketplace H2IOSC.

## 5. IL TRAINING NEL MARKETPLACE H2IOSC

Il *Marketplace* H2IOSC è una piattaforma progettata con l'obiettivo di offrire un punto di accesso centralizzato a dati, strumenti e servizi relativi alle scienze umane, sociali e culturali (Sichera, 2024). Consente l'integrazione e l'aggregazione di risorse interne ed esterne grazie a protocolli standard e API dedicate che permettono agli utenti di condividere, tra le altre risorse, anche materiali formativi. La sua struttura si ispira al *Social Sciences and Humanities Marketplace*<sup>5</sup>, con cui condivide una tipologia simile di oggetti (Barbot, 2024).

In un tale contesto, il *Marketplace* di H2IOSC accoglierà, insieme ad altre tipologie di dati, i metadati con cui sono state descritte le risorse didattiche sulle due piattaforme, rispettivamente *H2IOSC Training Environment* per i **corsi** diretti agli studenti e *H2IOSC Training Library* per i **materiali didattici** riutilizzabili da docenti e formatori. Tali metadati saranno allineati con quelli presenti nel modello dati del *Marketplace* anche se, naturalmente, le rispettive piattaforme mostreranno una metadatazione a grana più fine rispetto a quella del *Marketplace*, che si configura come una *discovery platform*. Inoltre, per quanto riguarda i vocabolari controllati associati a specifici metadati, questi saranno tratti dai vocabolari dello SSHOC *Marketplace*, con eventuali traduzioni in italiano.

Nel *Marketplace* saranno integrate funzionalità di ricerca tra oggetti di tipo diverso, per cui l'utente avrà a disposizione un'interfaccia di che potrà restituire, sulla base di un'interrogazione con chiave di ricerca, diverse tipologie di oggetti, comprendendo i materiali didattici di *H2IOSC Training Library* (suddivisi in corsi, sezioni, unità, moduli e *learning object*) ed eventi formativi o corsi in autoformazione disponibili su *H2IOSC Training Environment*, che saranno integrati nel *Marketplace* attraverso un calendario. Il *Marketplace* inoltre supporta la creazione di *workflow* complessi, in cui i materiali didattici possono essere aggregati in sequenze di apprendimento. Questo è particolarmente utile per formatori che desiderano integrare più risorse in un unico percorso formativo e garantirà l'interoperabilità con altre infrastrutture di *training*, come DARIAH-Campus,<sup>6</sup> CLARIN Learning Hub<sup>7</sup> e SSHOC Training Discovery Toolkit.<sup>8</sup>

In questo modo, il *Marketplace* H2IOSC permetterà alle piattaforme di training di creare un ecosistema integrato e dinamico, in grado di supportare l'apprendimento continuo e la condivisione delle conoscenze nel contesto delle Scienze Umane, Sociali e Culturali.

## 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 abstract.

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.

## BIBLIOGRAFIA

Barbot, L., Dolinar, M., Gray, E. J., Grisot, C., Illmayer, K., Kurzmeier, M., & McGillivray, B. (2024). Contextualizing Research Tools & Services Through Workflows in the SSH Open Marketplace. *Journal of Open Humanities Data*, 10(1). <https://doi.org/10.5334/johd.192>

<sup>5</sup> <https://marketplace.sshopencloud.eu/>

<sup>6</sup> <https://campus.dariah.eu/>

<sup>7</sup> <https://www.clarin.eu/content/learning-hub>

<sup>8</sup> <https://sshopencloud.eu/ssh-training-discovery-toolkit>

- Borek, L., Hastik, C., Khramova, V., & Geiger, J. (s.d.). *TaDiRAH Taxonomy of Digital Research Activities in the Humanities*. Social Sciences & Humanities Open Marketplace. Recuperato 21 gennaio 2025, da <https://marketplace.sshopencloud.eu/tool-or-service/w6OC9p>
- 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>.
- Filiposka, S., Green, D., Mishev, A., Kjorveziroski, V., Corleto, A., Napolitano, E., Paolini, G., Di Giorgio, S., Janik, J., Schirru, L., Gingold, A., Hadrosseck, C., Souyiultzoglou, I., Leister, C., Pavone, G., Sharma, S., Mendez Rodriguez, E. M., & Lazzeri, E. (2023). *D2.2 Methodology for FAIR-by-Design Training Materials*. <https://zenodo.org/records/8305540>
- Frontini, F., Gamba, F., Monachini, M., & Broeder, D. (2021). *SSHOC Multilingual Data Stewardship Terminology* [Dataset]. <http://hdl.handle.net/20.500.11752/ILC-567>
- Hoebelheinrich, N. J., Biernacka, K., Brazas, M., Castro, L. J., Fiore, N., Hellström, M., Lazzeri, E., Leenarts, E., Martinez Lavanchy, P. M., Newbold, E., Nurnberger, A., Plomp, E., Vaira, L., van Gelder, C. W. G., & Whyte, A. (2022). *Recommendations for a minimal metadata set to aid harmonised discovery of learning resources*. <https://doi.org/10.15497/RDA00073>
- Pedonese, G., Frontini, F., Ottaviani, R., Boschetti, F., Spadi, A., Francalanci, L., Scognamiglio, A., Restaneo, P., Chaban, A., Striova, J., & Benassi, L. (2024). *Materiali didattici come oggetti digitali FAIR: Una metodologia condivisa per la formazione in H2IOSC*. <https://doi.org/10.6092/unibo/amsacta/7927>
- Sichera, P., Marras, C., & Pasini, E. (2024). *ORCHESTRAZIONE API PER WORKFLOW APPLICATIVI NELL'AMBITO DELLE DIGITAL HUMANITIES*. Zenodo. <https://doi.org/10.5281/zenodo.14187534>
- van der Lek, Iulianna; Fišer, Darja. (2023). *Introduction to Language Data: Standards and Repositories*. In *UPSKILLS Learning Content*. [https://upskillsproject.eu/project/standards\\_repositories/](https://upskillsproject.eu/project/standards_repositories/). CC BY 4.0.
- van der Lek, I., Fišer, D., Frontini, F., & Pedonese, G. (2024, agosto 31). *Introduzione ai Dati Linguistici: Standard e Archivi Digitali*. Zenodo. <https://doi.org/10.5281/zenodo.13911935>
- van der Lek, I., Frontini, F., Fiser, D., & König, A. (2025). *Making the CLARIN Training Materials FAIR-By-Design*. CLARIN Annual Conference (CLARIN2023), Leuven. Zenodo. <https://doi.org/10.5281/zenodo.14699078>