CNR | Dipartimento Scienze Umane e Sociali Patrimonio culturale

DHC Digital Humanities Center

CNR DSU

Arcipelago di competenze DH al CNR:
metodi, strumenti, reti

a cura di: **Comitato di indirizzo del DHC**

01 ott 2026

Policy Brief

Dipartimento Scienze Umane e Sociali, Patrimonio Culturale
Digital Humanities Center
Collana Digital Humanities Center Policy Brief
n. 1/2026

Coordinamento scientifico
Federico Boschetti (CNR-ILC), Lucia Francalanci (CNR-OVI), Natale Rampazzo (CNR-IRIDIS), Lisa Reggiani (CNR-IRFiS), Alessia Scognamiglio (CNR-IRFiS), Daria Spampinato (CNR-ISTC), Michela Tardella (CNR-IRFiS)

Curatori del numero
Comitato di indirizzo
Progetto grafico e copertina
Silvestro Caligiuri (CNR-IRFiS)

Contatti
CNR-DSU. Piazzale Aldo Moro, 7 – 00185 Roma. Tel. +39 06 49933328
Mail: dhc-policy-brief@cnr.it

ISSN: 3103-9367
DOI: 10.36134/DHC-2026-1

ottobre 2026

INDICE

Salvatore Capasso, Cristina Marras, Daria Spampinato
La collana "Digital Humanities Center Policy Brief" 1

Tommaso Agnoloni, Lucia Francalanci, Emilio M. Sanfilippo
Mappare le Digital Humanities al CNR DSU attraverso i dati della ricerca: pubblicazioni, autori e parole chiave 3

Angelo Mario Del Grosso
Un ponte FAIR per le Digital Humanities: l'integrazione tra OpenAIRE e DH-ATLAS al CNR 27

Cristina Marras, Alessia Scognamiglio
Nodi, reti e sinergie 34

La collana “Digital Humanities Center Policy Brief”

La collana *Digital Humanities Center Policy Brief* nasce come strumento collaborativo, multi e interdisciplinare del CNR DSU - Centro Digital Humanities[1], e ha come finalità principale quella di formulare policy e linee guida, di valorizzare competenze e metodologie specifiche, di riflettere criticamente e analizzare temi emergenti che riguardano l’umanistica digitale nella sua integrazione, impatto e rappresentazione nella ricerca del CNR.
Questa serie di pubblicazioni è innanzitutto uno strumento di condivisione delle esperienze maturate al CNR per rendere accessibili dati, servizi e risultati delle ricerche di Digital Humanities (DH) e per contribuire a una maggiore consapevolezza della forza sinergica prodotta dall’incontro delle scienze umane e sociali con le metodologie computazionali e gli strumenti digitali.
La programmazione editoriale della collana[2] prevede tematiche che il Comitato di indirizzo del Centro ha individuato come strategiche e sensibili per il CNR, tra le quali le metodologie, il diritto, l’etica, la formazione, le competenze, la valutazione della ricerca nel settore DH. Uno sguardo particolare si intende dare alla governance e pianificazione strategica per le DH e al loro impatto sociale, economico e civico; non ultimo l’uso e l’integrazione dell’Intelligenza Artificiale nelle pratiche DH (come Machine Learning e Large Language Models), tema di particolare attualità e urgenza per le scienze umane e sociali.
A partire dalla difficoltà di inquadrare e definire l’area di ricerca delle Digital Humanities, la condivisione delle esperienze maturate al CNR nell’ambito di questo settore variegato e composito potrà attivare e favorire l’innesto capillare e la reale integrazione del digitale e computazionale – in termini di modelli, metodi, processi – anche in contesti plurali ancora distanti dalla tecnologia.

Questo primo numero, come indicato dal titolo, riflette proprio sull’Arcipelago di competenze DH al CNR per focalizzarsi sui metodi, gli strumenti e le reti.
Il contributo “Mappare le Digital Humanities al CNR DSU attraverso i dati della ricerca: pubblicazioni, autori e parole chiave” dei colleghi Tommaso Agnoloni, Lucia Francalanci ed Emilio M. Sanfilippo presenta non solo dati di particolare rilevanza, aggregati per la prima volta, a coprire l’area di ricerca DH, ma offre anche un modello ragionato e replicabile per la loro raccolta e analisi.
Si tratta di una attività del Centro che si intende aperta, non solo in quanto i dati vanno periodicamente aggiornati, integrati e riaggregati, ma anche perché la complessità del lavoro richiede il contributo e la collaborazione di tutti. La mappatura deve coinvolgere infatti tutto il personale che opera in questo settore, sia sul piano della accuratezza dei dati della ricerca relativi alle DH che vengono esposti nelle diverse piattaforme istituzionali, sia sul piano più

[1] https://www.dsu.cnr.it/digital-humanities-center/ (ultima consultazione dei siti web: 5/10/2026).
[2] https://www.dsu.cnr.it/prodotti/tipologia/collane/digital-humanities-center-policy-brief/#descrizione-estesa

operativo-collaborativo con il gruppo mappatura che ha avviato questa importante attività per il Centro.

La seconda sezione, “Un ponte FAIR per le Digital Humanities: l’integrazione tra OpenAIRE e DH-ATLAS al CNR” di Angelo Mario Del Grosso, risponde alla necessità di offrire alla comunità scientifica approcci e strumenti capaci di integrare la granularità disciplinare e l’interoperabilità per rendere la ricerca DH reperibile, interrogabile e riusabile. Si concentra, anche attraverso la descrizione di alcuni esempi, sugli strumenti per la ricostruzione della rete di relazioni tra persone, temi, progetti, metodi e servizi, possibile grazie alla conformità dei dati ai principi FAIR (*Findable*, *Accessible*, *Interoperable*, *Reusable*). L’accento dunque cade inevitabilmente sulle comunità che questi dati li producono, li gestiscono e li preservano.
Il fascicolo si chiude con una breve riflessione su “Nodi, reti e sinergie”, in cui le autrici Alessia Scognamiglio e Cristina Marras propongono un framework strategico, sostenuto dai dati della mappatura, che si articola su sette pilastri programmatici. La capacità interna, fatta di competenze, metodologie e sfide della ricerca di frontiera, è messa in relazione con l’interoperabilità sistemica data da alcuni asset portanti, quali per esempio le infrastrutture e la scienza aperta; tenere conto di questi pilastri sostiene la ricerca DH del CNR e ne consolida la presenza nei network globali della conoscenza oltreché massimizzarne l’impatto scientifico.

Nel concludere questa premessa al primo numero, inaugurale, della collana *Digital Humanities Center Policy Brief* vogliamo sottolineare ancora una volta l’intento e l’obiettivo sotteso a questo progetto editoriale: aprire un cantiere di lavoro e di scambio a beneficio di tutto il personale e gli organismi dell’Ente.

Per il lavoro di preparazione del numero desideriamo ringraziare le colleghe Michela Tardella e Lisa Reggiani. Per la mappatura dei dati DH al CNR ringraziamo per la disponibilità e la collaborazione le colleghe e i colleghi: Leonardo Pica Ciamarra, referente per l’area disciplinare SH_5; Emma Lazzeri, Responsabile Sezione Open Science del CNR; Alessia Bardi, rappresentante CNR all’OpenAIRE General Assembly; Riccardo Del Gratta e Giorgia Rubin, Responsabile Scientifico Unità CNR e ricercatrice del progetto ATLAS. Questo progetto editoriale è possibile grazie alla disponibilità e alla collaborazione dello staff del DSU, in particolare di Giulia Antonini e Stefano Carotenuto.

Salvatore Capasso, Cristina Marras, Daria Spampinato

Mappare le Digital Humanities al CNR DSU attraverso i dati della ricerca: pubblicazioni, autori e parole chiave

Mapping the Digital Humanities at the CNR DSU using Research Data: Publications, Authors and Keywords

Tommaso Agnoloni, CNR-IGSG, tommaso.agnoloni@cnr.it
Lucia Francalanci, CNR-OVI, lucia.francalanci@cnr.it
Emilio M. Sanfilippo, CNR-ISTC, emiliomaria.sanfilippo@cnr.it

Abstract
Il contributo presenta una prima attività di mappatura delle competenze e delle linee di ricerca riconducibili alle Digital Humanities all'interno del CNR. L'indagine si basa sull'analisi dei metadati e delle parole chiave delle pubblicazioni presenti nel repository istituzionale CNR-IRIS, successivamente normalizzate e organizzate in una tassonomia, e su un questionario rivolto al personale di ricerca del DSU.

Abstract (English)
This paper presents an initial attempt to map the expertise and research areas related to Digital Humanities within the CNR. The study is based on an analysis of the metadata and keywords of publications in the CNR-IRIS institutional repository, which were subsequently standardized and organized into a taxonomy, and on a survey of research staff at the DSU.

Keywords
digital humanities – mapping – metadata – research data – taxonomy

Introduzione

Le Digital Humanities (DH) costituiscono uno dei principali ambiti di convergenza tra discipline umanistiche, scienze sociali e tecnologie dell'informazione. La loro natura trasversale rende particolarmente interessante comprendere come tali competenze siano distribuite all'interno di un Ente di Ricerca multidisciplinare come il Consiglio Nazionale delle Ricerche.

L'obiettivo di questo lavoro non è quello di "misurare" la produzione scientifica, bensì delineare un quadro preliminare sulla ricerca DH. L'indagine mira in particolare a far emergere la natura fortemente interdisciplinare di questa comunità, distribuita tra diversi istituti e caratterizzata da un'ampia varietà di interessi scientifici, che spaziano dalla linguistica computazionale all'archeologia digitale, all'intelligenza artificiale, alla modellazione della conoscenza, fino all'informatica giuridica e alla valorizzazione del patrimonio culturale.

Il contributo si articola in tre sezioni: la prima descrive l'individuazione degli autori e l'estrazione dei dati relativi alle pubblicazioni scientifiche; la seconda presenta la costruzione e l'analisi di

una tassonomia delle parole chiave; la terza amplia l'indagine attraverso i risultati di un questionario rivolto al personale di ricerca del CNR DSU.

1. Selezione delle fonti di dati e costruzione del criterio di filtraggio

Fra le fonti di dati esistenti che raccolgono i metadati bibliografici delle pubblicazioni scientifiche (*scholarly graphs*, come OpenAIRE[1], OpenAlex[2] e alternative commerciali) non ne esiste alcuna già strutturata e sufficientemente accurata che consenta di individuare in modo diretto i ricercatori CNR attivi nell'ambito delle Digital Humanities, i prodotti della loro ricerca e i temi trattati attraverso le parole chiave associate alle pubblicazioni.

La fonte primaria scelta nel nostro caso è il repository istituzionale CNR-IRIS, che raccoglie i metadati di tutti i prodotti della ricerca depositati dai ricercatori CNR ed è accessibile tramite API[3]. CNR-IRIS non dispone però di tag tematici espliciti che permettano di filtrare direttamente i record di interesse per le DH: in assenza di tale chiave, si è scelto di utilizzare come criterio di accesso liste precostituite di ricercatori CNR il cui profilo di ricerca includesse, con ragionevole approssimazione, il campo delle Digital Humanities.

Costruzione delle liste di autori

Sono state sperimentate due opzioni per compilare le liste di partenza:

1. **Lista SH5_11** — 94 ricercatori e tecnologi CNR che nell'ultima mappatura istituzionale dell'afferenza alle aree disciplinari (2025) avevano dichiarato fra i propri settori di interesse l'ambito SH5_11 (*Digital humanities; computational and digital approaches in the cultural sphere*).
2. **Lista AIUCD** — 136 ricercatori e tecnologi CNR identificati come autori o co-autori di contributi presentati alle edizioni della conferenza AIUCD[4] negli ultimi dieci anni (2016–2025).

Le due liste sono state curate e arricchite con identificatori univoci standard — ORCID e rp_id (identificativo interno al sistema CNR-IRIS) — così da poter accedere in modo univoco ai record, selezionando solo i contributi dei quali i ricercatori risultano autori.

Le liste si sovrappongono parzialmente (Figura 1): 40 ricercatori compaiono in entrambe. La sovrapposizione corrisponde al 42,6% della lista SH5_11 e al 29,4% della lista AIUCD, a conferma del fatto che i due criteri di selezione intercettano popolazioni in larga parte diverse: 54 ricercatori sono presenti solo in SH5_11 e 96 solo in AIUCD.

[1] https://explore.openaire.eu (ultima consultazione dei siti web: 5/10/2026).

[2] https://openalex.org

[3] https://iris.cnr.it/

[4] Associazione per l'Informatica Umanistica e la Cultura Digitale: https://www.aiucd.it

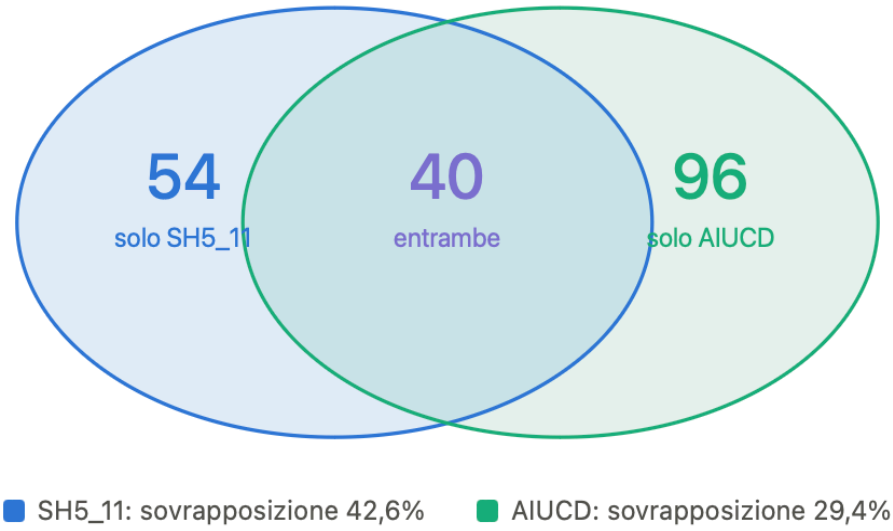

Figura 1. Numerosità e sovrapposizione delle liste di autori di partenza: SH5_11 e AIUCD.

Distribuzione per istituto. La distribuzione per istituti degli autori CNR è illustrata nelle Figure 2 e 3. CNR-ILC è l'Istituto predominante in entrambe le liste, ma con un peso proporzionalmente molto maggiore in AIUCD (36 vs. 23). Le divergenze più rilevanti riguardano CNR-ISTI (3 in SH5_11, 20 in AIUCD), CNR-IMATI (3 vs. 10) e CNR-ISPF (assente in SH5_11, 6 in AIUCD). In direzione opposta, CNR-OVI è più rappresentato in SH5_11 (16 vs. 9) e CNR-ITD è presente solo in SH5_11 (6 autori, assente in AIUCD).

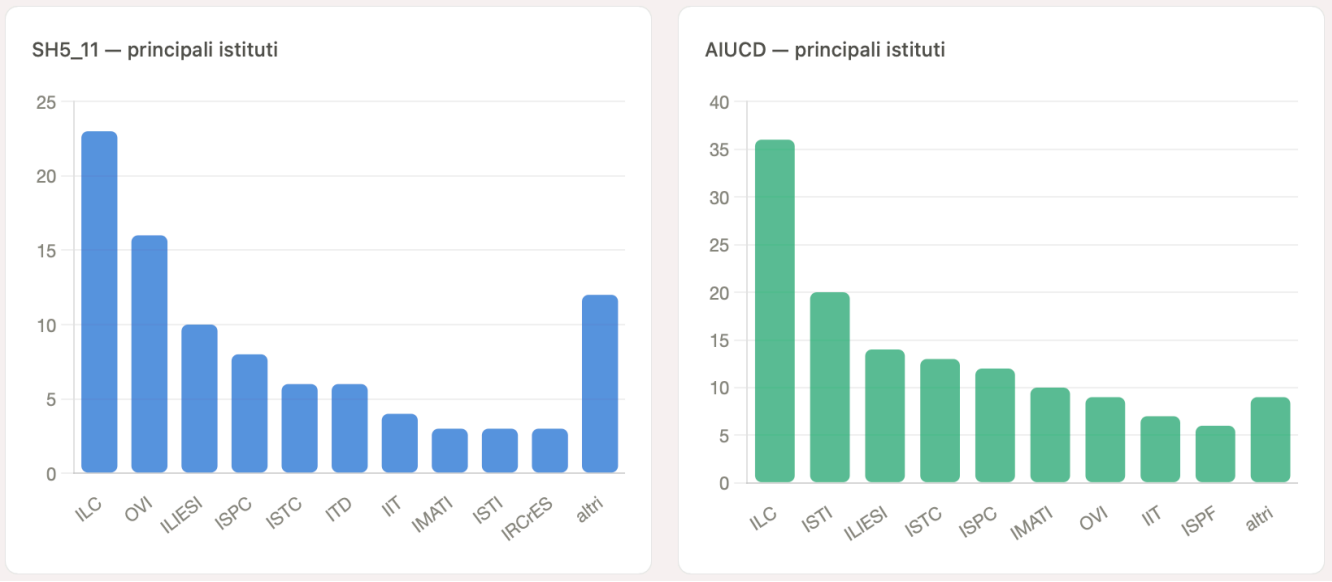

Figura 2. Distribuzione per istituto.

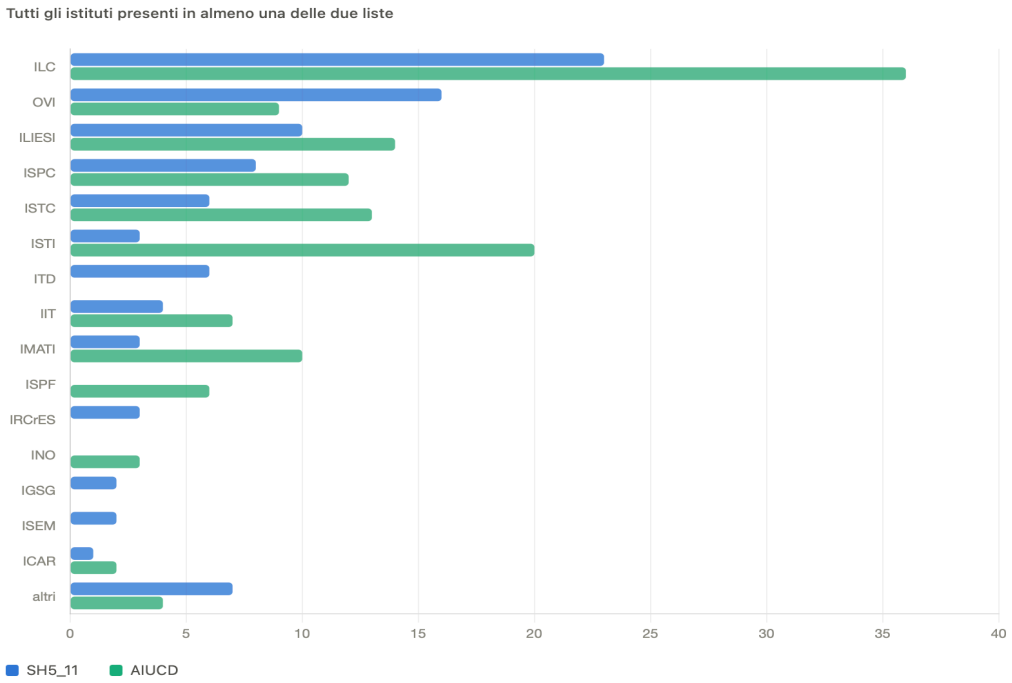

Figura 3. Confronto per istituto: SH5_11 vs. AIUCD.

Composizione per dipartimento. La Figura 4 riporta la distribuzione degli autori per dipartimento CNR. Entrambe le liste sono a netta prevalenza DSU, ma AIUCD presenta una componente DIITET (Dipartimento di Ingegneria, Informatica e Tecnologie per l'Energia e i Trasporti) molto più ampia (39 vs. 12), riflesso della forte presenza di ISTI, IMATI e istituti analoghi nei *proceedings* della conferenza. La lista AIUCD include inoltre un piccolo numero di autori afferenti ai dipartimenti di scienze biomediche e scienze fisiche, del tutto assenti in SH5_11.

Dipartimento	SH5_11	AIUCD
Scienze umane e sociali, patrimonio culturale (DSU)	77	91
Ingegneria, ICT e tecnologie per l'energia e i trasporti (DIITET)	12	39
Scienze biomediche	—	3
Scienze fisiche e tecnologie della materia	—	3
Amministrazione centrale	3	—
Scienze del sistema terra e tecnologie per l'ambiente	2	—

Figura 4. Distribuzione per dipartimento CNR.

Nella lista SH5_11, il settore SH5_11 è dichiarato come primo settore da 35 ricercatori (37,2%), come secondo da 34 (36,2%) e come terzo da 25 (26,6%). La distribuzione (Figura 5) è

sorprendentemente uniforme, il che evidenzia che SH5_11 non costituisce necessariamente la specializzazione primaria del personale che lo dichiara: poco meno dei due terzi lo colloca in seconda o terza posizione.

CNR-ILC è l'Istituto con la maggiore concentrazione di ricercatori e tecnologi per i quali SH5_11 è il settore principale (14 su 35), seguito da CNR-OVI (4) e CNR-IIT (3). CNR-ILIESI e CNR-OVI presentano una quota significativa di personale che dichiara SH5_11 in seconda posizione, a indicare un profilo disciplinare più tradizionalmente orientato alle scienze umane con una componente digitale secondaria.

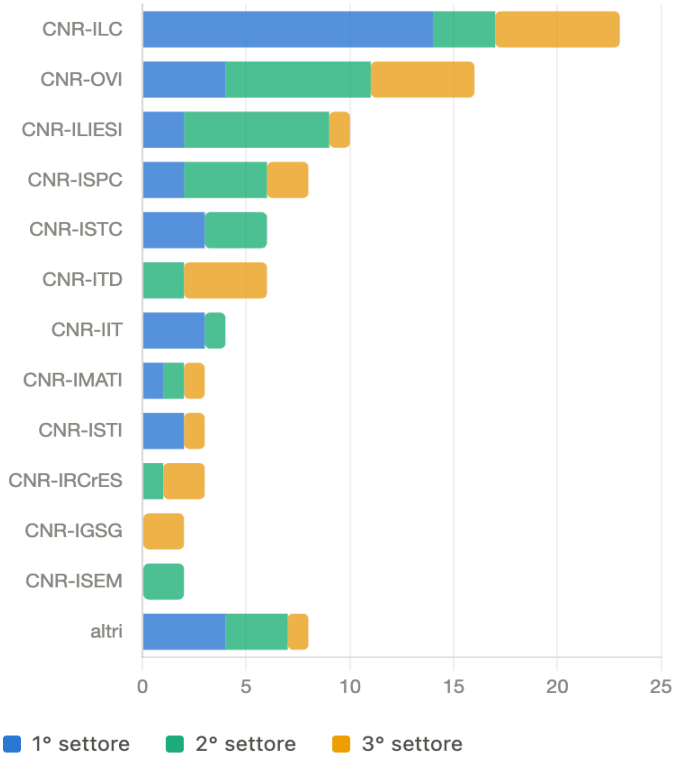

Figura 5. Autori per istituto, suddivisi per posizione di SH5_11 (1°, 2°, 3°).

Quando SH5_11 è il settore principale, i settori più frequentemente co-dichiarati sono PE6_10 (*Web and information systems, data management systems, information retrieval and digital libraries, data fusion*, 14 occorrenze), SH4_9 (*Theoretical linguistics; computational linguistics*, 10) e PE6_7 (*Artificial intelligence, intelligent systems, natural language processing*, 9). Quando SH5_11 è in seconda o terza posizione, il settore principale è prevalentemente SH5_4 (*Philology; text and image studies*) o SH3_9 *(Technological, methodological and social aspects of teaching and learning, curriculum studies, education and educational policies*) / SH4_10 (*Language typology; historical linguistics*), indicando una provenienza più radicata nelle scienze umane tradizionali.

Esiste una correlazione forte tra la posizione di SH5_11 e la presenza nella lista AIUCD (Figura 6): il 65,7% degli autori con SH5_11 come settore principale compare anche in AIUCD, contro

il 38,2% per il secondo settore e appena il 16,0% per il terzo. Questo suggerisce che chi dichiara SH5_11 come settore principale è molto più attivo in senso stretto nelle DH, con una partecipazione effettiva alle conferenze di settore.

Posizione SH5_11	Autori	Anche in AIUCD	Solo SH5_11	% sovrapposizione
1° settore	35	23	12	65,7%
2° settore	34	13	21	38,2%
3° settore	25	4	21	16,0%
Totale	94	40	54	42,6%

Figura 6. Posizione SH5_11 vs. AIUCD.

Scelta del criterio di filtraggio

Entrambe le liste presentano limiti intrinseci: si tratta in ogni caso di una selezione *"sporca"*, che richiede una valutazione qualitativa di pertinenza sulle keyword estratte e sulle liste stesse. La lista SH5_11, in particolare, include ricercatori e tecnologi i cui prodotti non sono necessariamente riconducibili alle DH, e può al contempo escludere autori attivi nel campo ma non afferenti a quell'area disciplinare. La lista AIUCD, per contro, si basa su una fonte più diretta, la partecipazione a una conferenza di settore, e ha dato risultati qualitativamente più omogenei.

Il criterio di filtraggio finale adottato per l'estrazione dei record CNR-IRIS è pertanto quello basato sulla lista degli autori ottenuta dallo spoglio dei *proceedings* AIUCD.

Estrazione dei dati da IRIS

A partire dagli identificatori degli autori selezionati, sono stati estratti tramite le API di CNR-IRIS i record delle pubblicazioni nelle seguenti tipologie: 01.01 Articolo in rivista, 02.01 Contributo in volume (Capitolo o Saggio), 04.01 Contributo in Atti di convegno.

Come riportato nella Tabella 1, risultano 4.013 record che coprono il periodo 1969–2026[5].

Tipo	**Numero**	**Anni**
Articolo in rivista	1.496	1969–2026
Contributo in volume (Capitolo o Saggio)	536	1985–2026
Contributo in Atti di convegno	1.981	1970–2026
Totale	4.013	

Tabella 1. Numero di record e copertura temporale per tipologia di contributo.

[5] Estrazione dei dati da CNR-IRIS effettuata a maggio 2026.

Il corpus cresce in modo significativo a partire dal 2003, con i contributi in atti di convegno che si confermano il tipo dominante nel corso degli anni (Figura 7).

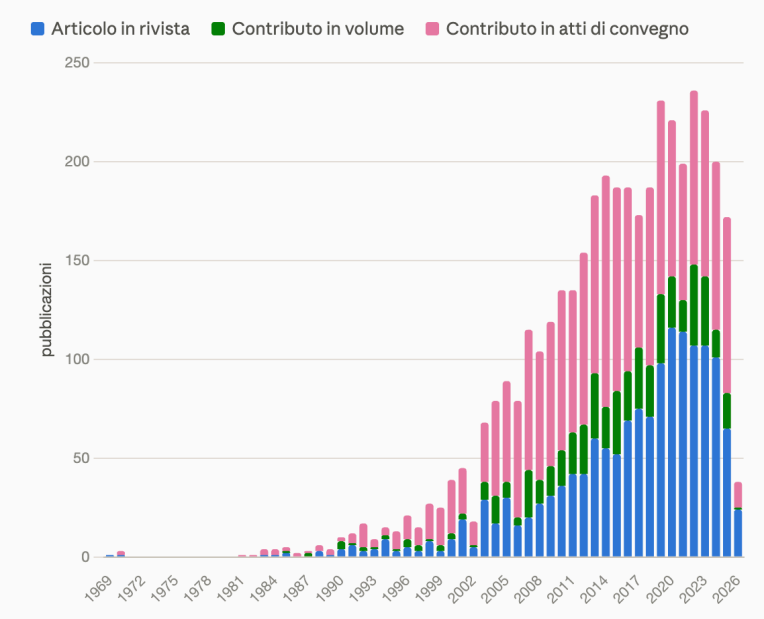

Figura 7. Distribuzione temporale dei record per tipologia di contributo.

Le parole chiave associate ai record non sono direttamente accessibili tramite API, ma lo sono indirettamente tramite *scraping* delle pagine HTML associate ai prodotti. Esse sono inoltre presenti indifferentemente in italiano e in inglese, spesso senza indicazione esplicita della lingua.

La lista delle parole chiave associate ai record estratti include le parole chiave presenti nei record IRIS dopo una prima elaborazione automatica di normalizzazione sintattica, insieme alla loro frequenza di occorrenza e ai record di provenienza costituisce la base di partenza per l'ispezione e il raffinamento manuale. La lista completa risultante include 6.509 parole chiave.

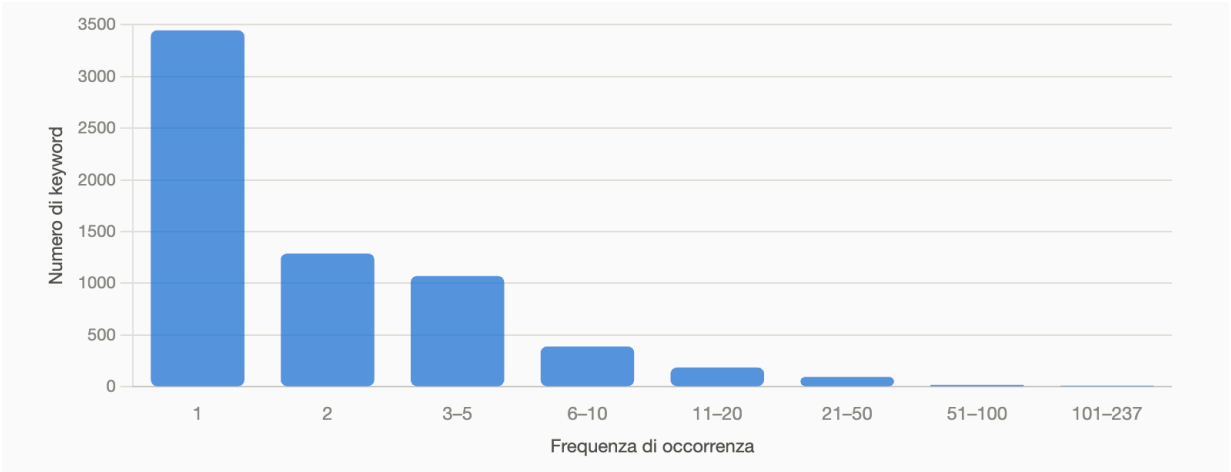

Figura 8. Numero di parole chiave per classe di frequenza.

Come mostrato nella Figura 8, la distribuzione è fortemente asimmetrica con coda lunga, andamento tipico dei vocabolari di testo libero. Più della metà delle 6.509 parole chiave (52,9%) occorre una sola volta nell'intero corpus (hapax), e quasi tre quarti (72,7%) compaiono al massimo due volte. La curva scende rapidamente: 1.071 parole chiave appaiono 3–5 volte, 390 tra 6 e 10 volte, e solo 315 (4,8%) raggiungono frequenza maggiore di 10. Appena 12 parole chiave superano le 100 occorrenze.

La forte concentrazione di termini a bassissima frequenza riflette l'assenza di un vocabolario controllato condiviso: le parole chiave accessibili da IRIS sono state inserite liberamente dai ricercatori senza un sistema di normalizzazione, il che produce varianti grafiche, sinonimi, traduzioni e scelte terminologiche personali che disperdono il contenuto informativo in migliaia di voci di fatto equivalenti ma non riconciliate.

Questo pattern ha implicazioni dirette per la fase di pulizia: una soglia minima di frequenza > 1 ridurrebbe il vocabolario della metà eliminando una buona parte del rumore pur conservando la copertura tematica sostanziale.

Un'ulteriore fonte di complessità è la natura multilingue del vocabolario: le parole chiave sono presenti sia in italiano sia in inglese, spesso in modo non dichiarato. Solo 946 delle 6.509 parole chiave riporta un tag linguistico esplicito (221 in italiano, 725 in inglese); per le restanti 5.563 si è tentato di attribuire la lingua automaticamente tramite librerie di *language detection* e regole euristiche (1.070 italiano, 4.099 inglese, 394 non attribuita). Data la netta prevalenza dell'inglese, si è scelto in questa prima fase di adottarla come lingua di riferimento, traducendo automaticamente le parole chiave italiane e applicando successivamente una deduplicazione.

La lista grezza delle parole chiave così ottenuta costituisce un semilavorato da analizzare, normalizzare e ripulire per ottenere una risorsa concettuale gestibile dei temi DH al CNR.

2. Costruzione di una tassonomia delle parole chiave per il dominio delle Digital Humanities

La lista grezza delle parole chiave ottenuta dall'estrazione costituisce una risorsa ricca ma fortemente eterogenea e non strutturata. Dal momento che un semplice conteggio delle occorrenze non consentirebbe un'analisi precisa ed esaustiva dei temi della ricerca, si è reso necessario definire una procedura di organizzazione del vocabolario, volta a ridurre la variabilità formale delle parole chiave e a ricondurle a una struttura coerente. Il processo, che combina tecniche automatiche e revisione manuale, si è articolato in tre fasi principali: normalizzazione terminologica, classificazione semantica e validazione finale.

Normalizzazione terminologica

La normalizzazione è finalizzata a ridurre la variabilità formale delle parole chiave, in modo da garantire coerenza terminologica e confrontabilità dei dati nella successiva fase di classificazione.

L'attività è consistita principalmente nel ricondurre a una forma standard parole chiave equivalenti ma formalmente o linguisticamente differenti, preservando per quanto possibile la

tracciabilità tra i termini originali e quelli normalizzati. Come lingua di riferimento è stato scelto l'inglese, in quanto prevalente nella comunicazione scientifica internazionale e nelle principali tassonomie del settore. L'adozione della lingua inglese favorisce inoltre l'interoperabilità con risorse esterne.

Una prima elaborazione è stata effettuata mediante il supporto di un Large Language Model (LLM); in particolare, il prompt è stato eseguito con Claude OPUS 4.8. I risultati sono stati successivamente verificati e corretti manualmente.

La normalizzazione ha riguardato, oltre alla traduzione delle parole chiave italiane nelle corrispondenti forme inglesi, l'uniformazione di varianti ortografiche e morfologiche, la gestione degli acronimi e delle relative forme estese, la correzione di errori tipografici, la pulizia dei caratteri e l'eliminazione dei duplicati.

Per ciascuna parola chiave sono state registrate sia la forma standard, sia le eventuali varianti. Le forme sono state unificate solo in presenza di un'equivalenza certa.

I principali criteri adottati sono riportati nella Tabella 2.

Criterio	**Descrizione**
Uniformazione linguistica	Traduzione delle parole chiave italiane nella forma inglese maggiormente diffusa nel settore di riferimento; in presenza di più traducenti possibili, scelta della denominazione prevalente nella comunicazione scientifica internazionale (*Patrimonio culturale* → *Cultural Heritage*).
Varianti ortografiche e morfologiche	Riconduzione delle grafie alternative a un'unica forma, unificazione di singolari e plurali e differenze di capitalizzazione (*digitisation* → *Digitization*; *ontology/ontologies* → *Ontology*).
Acronimi	Normalizzazione degli acronimi in maiuscolo e riconduzione delle formulazioni estese all'acronimo, quando quest'ultimo costituisce la forma d'uso consolidata nella comunità scientifica (*Natural Language Processing* → *NLP*).
Capitalizzazione	Adozione del *Title Case* per le locuzioni e del maiuscolo per gli acronimi, nel rispetto delle eventuali convenzioni ufficiali (*machine learning* → *Machine Learning*; *nlp* → *NLP*).
Pulizia formale	Eliminazione di spazi iniziali e finali, spazi multipli, punteggiatura non significativa, caratteri speciali superflui e separatori incoerenti (*A.I.*, *A-I*, *A I* → *AI*).
Deduplicazione	Aggregazione delle occorrenze riconducibili alla medesima forma canonica, preservando il collegamento con ciascun termine originale.
Denominazioni ufficiali	Conservazione della denominazione ufficiale di software, linguaggi di programmazione, standard, prodotti e marchi; i nomi commerciali non sono stati tradotti né le loro sigle espanse automaticamente (*Python*, *TEI*).

Criterio	Descrizione
Nomi propri	Conservazione della denominazione originale di antroponimi, toponimi, enti e progetti, intervenendo soltanto sulla correzione di errori o sull'uniformazione di varianti puramente grafiche.

Tabella 2. Criteri di normalizzazione delle parole chiave.

Classificazione semantica delle parole chiave

L'applicazione dei criteri di normalizzazione ha prodotto un vocabolario costituito da 5.468 parole chiave univoche. Sebbene la normalizzazione abbia ridotto la variabilità formale del lessico, risultava necessaria un'ulteriore fase di organizzazione semantica delle parole chiave. È stata quindi definita una tassonomia articolata in dieci categorie (Tabella 3), individuate attraverso un'analisi esplorativa del vocabolario normalizzato.

La classificazione è stata effettuata in modo automatico mediante un LLM istruito sulla base delle descrizioni delle categorie individuate (il prompt è stato eseguito con Claude OPUS 4.8). Per ciascuna parola chiave il modello ha proposto l'assegnazione a un'unica categoria, unitamente a un indice di confidenza e, nei casi di maggiore incertezza, a un breve commento esplicativo.

Categoria	Descrizione	Esempi
Antroponimi	Nomi di persone, studiosi o personalità storiche.	*Giambattista Vico, Immanuel Kant, Vincenzo Bellini*
Attività	Processi, operazioni e pratiche svolte nell'ambito della ricerca.	*Annotation, Digitization, Transcription, 3D Modelling*
Concetti/Temi	Concetti e temi che identificano gli argomenti trattati dalla ricerca.	*Accessibility, Cultural Heritage, Neologisms, Open Access*
Dati/Oggetti digitali	Risorse, materiali e tipologie di dati utilizzati, prodotti o analizzati nella ricerca.	*Text Corpora, Manuscripts, Digital Editions, Dataset*
Discipline/Ambiti disciplinari	Aree disciplinari o settori di studio che identificano il dominio della ricerca.	*Archaeology, Linguistics, Philology*
Istituzioni/Infrastrutture di ricerca	Enti, istituzioni, comunità scientifiche e infrastrutture coinvolti nella ricerca.	*AIUCD, CLARIN, DARIAH, EOSC*
Metodologie/Approcci	Metodologie, tecniche e approcci di analisi e ricerca.	*NLP, Stylometry, Computer Vision*

Categoria	**Descrizione**	**Esempi**
Progetti	Progetti di ricerca, iniziative identificate da un nome proprio.	*ARCO, H2IOSC, Bellini Digital Correspondence*
Tecnologie/Standard/ Strumenti	Tecnologie, linguaggi, standard, software e strumenti utilizzati nelle attività di ricerca.	*XML, TEI, OCR, API, AI*
Toponimi	Nomi geografici riferiti a luoghi, città o territori.	*Florence, Rome, Sicily*

Tabella 3. Categorie della tassonomia delle parole chiave.

Sono state classificate unicamente le parole chiave con frequenza pari o superiore a due, escludendo, pertanto, le forme che occorrono una sola volta (hapax). Tale scelta è motivata dalla distribuzione fortemente asimmetrica del vocabolario, caratterizzato da un'elevata presenza di termini a bassissima frequenza. L'applicazione di una soglia minima di frequenza consente di ridurre sensibilmente il rumore, pur preservando la copertura delle principali aree tematiche del corpus. Il dataset sottoposto a classificazione è perciò costituito da 2.828 parole chiave normalizzate.

Revisione della classificazione

La classificazione automatica è stata successivamente sottoposta a una revisione manuale da parte di un gruppo di lavoro interdisciplinare costituito da sei ricercatori e tecnologi. L'obiettivo era valutare la correttezza delle assegnazioni proposte dal modello, risolvere i casi ambigui e verificare la coerenza della classificazione. Tale revisione era particolarmente rilevante anche alla luce dell'elevato numero di assegnazioni a bassa confidenza prodotte dal modello (1.144 parole chiave, pari al 40,4% del totale).

Per supportare la disambiguazione delle parole chiave e verificare la coerenza dello schema di classificazione, i revisori hanno fatto riferimento ad alcune tassonomie e vocabolari già consolidati nella comunità delle DH, utilizzati esclusivamente come supporto interpretativo (Tabella 4).

Categoria	**Tassonomie e vocabolari**
Attività	• TaDiRAH
Concetti/Temi	• Nuovo soggettario BNCF
Dati/Oggetti digitali	• KNOT Tassonomia • ATLAS

Categoria	**Tassonomie e vocabolari**
Discipline/Ambiti disciplinari	• DHA Taxonomy
Istituzioni/Infrastrutture di ricerca	• ROR [per normalizzazione] • DHA Taxonomy
Metodologie/Approcci	• TaDiRAH • EOSC Resource Category
Tecnologie/Standard/Strumenti	• DHA Taxonomy • SSK Standard • EOSC Resource Category

Tabella 4. Tassonomie e vocabolari di riferimento.

In questa fase sono emerse alcune criticità, quali la presenza di concetti a diversi livelli di generalità, la polisemia di alcuni termini, difficilmente interpretabili se estrapolati dal contesto di riferimento, e la presenza di parole chiave potenzialmente riconducibili a più categorie.
Si è scelto, infatti, di attribuire a ciascun termine un'unica categoria prevalente, così da garantire coerenza e confrontabilità nell'analisi quantitativa: tale scelta ha comportato un accurato lavoro di rifinitura, basato sulla valutazione dei contesti d'uso e sull'analisi del significato dei singoli termini, per individuare la categoria più appropriata.

Analisi dei risultati

A seguito della revisione manuale, che ha comportato la riclassificazione di alcuni termini e l'accorpamento di parole chiave semanticamente equivalenti, il vocabolario finale risulta costituito da 1.494 parole chiave, distribuite nelle dieci categorie della tassonomia (Tabella 5).

Categoria	**N. parole chiave**	**%**
Dati/Oggetti digitali	356	23,83%
Metodologie/Approcci	341	22,82%
Concetti/Temi	196	13,12%
Tecnologie/Standard/Strumenti	179	11,98%
Attività	174	11,65%
Discipline/Ambiti	167	11,18%
Antroponimi	24	1,61%
Istituzioni/Infrastrutture	22	1,47%
Toponimi	19	1,27%

Categoria	**N. parole chiave**	**%**
Progetti	16	1,07%
Totale	**1.494**	**100%**

Tabella 5. Distribuzione delle parole chiave per categoria.

Le categorie più rappresentate risultano Dati/Oggetti digitali (23,8%) e Metodologie/Approcci (22,8%), seguite da Concetti/Temi (13,1%), Tecnologie/Standard/Strumenti (12,0%), Attività (11,7%) e Discipline/Ambiti (11,2%). Le restanti categorie incidono ciascuna per meno del 2% del totale.
Le tabelle seguenti riportano, rispettivamente, le tre parole chiave più frequenti per ciascuna categoria della tassonomia e le parole chiave con frequenza assoluta più elevata.

Categoria	**1ª par. chiave**	**Occ.**	**2ª par. chiave**	**Occ.**	**3ª par. chiave**	**Occ.**
Attività	Data Visualization	142	Training	73	Digital Scholarly Editing	58
Antroponimi	Giambattista Vico	76	Gottfried Wilhelm von Leibniz	32	Vincenzo Bellini	25
Concetti/Temi	Accessibility	247	Cultural Heritage	226	FAIR	158
Dati/Oggetti digitali	Ontology	361	Digital Libraries	253	Linked Data	110
Discipline/Ambiti	Archaeology	297	Digital Humanities	199	Lexicography	124
Istituzioni/Infrastrutture	Research Infrastructures	137	CLARIN	70	OpenAIRE	23
Metodologie/Approcci	NLP	229	Information Retrieval	156	Knowledge Extraction	94
Progetti	H2IOSC	38	Bellini Digital Correspondence	23	VLO	11
Tecnologie/Standard/ Strumenti	Semantic Web	290	LOD	229	AI	138
Toponimi	Arabia	40	Sicily	25	Crete	16

Tabella 6. Parole chiave più frequenti per categoria.

Rango	Parola chiave	Occorrenze	Categoria
1	Ontology	361	Dati/Oggetti digitali
2	Archaeology	297	Discipline/Ambiti disciplinari
3	Semantic Web	290	Tecnologie/Standard/Strumenti
4	Digital Libraries	253	Dati/Oggetti digitali
5	Accessibility	247	Concetti/Temi
6	LOD	229	Tecnologie/Standard/Strumenti
7	NLP	229	Metodologie/Approcci
8	Cultural Heritage	226	Concetti/Temi
9	Digital Humanities	199	Discipline/Ambiti disciplinari
10	FAIR	158	Concetti/Temi

Tabella 7. Parole chiave a maggiore frequenza.

La parola chiave più frequente è *Ontology* (361 occorrenze), seguita da *Archaeology* (297) e *Semantic Web* (290), evidenziando la centralità delle tecnologie semantiche e della valorizzazione del patrimonio culturale nella ricerca svolta dagli istituti del CNR nel campo delle Digital Humanities. Risulta inoltre evidente la crescente attenzione verso la condivisione, l'accessibilità e il riuso delle risorse digitali, come attestato dalle occorrenze di *Accessibility* e *FAIR*.

Per approfondire ulteriormente la distribuzione dei termini più ricorrenti, è stata analizzata la loro ripartizione nelle diverse categorie della tassonomia (Tabella 8). Sono considerate le parole chiave con frequenza pari o superiore a 100, che risultano essere 21.

Categoria	N. parole chiave ≥100	%
Concetti/Temi	6	28,6%
Discipline/Ambiti disciplinari	4	19,0%
Tecnologie/Standard/Strumenti	4	19,0%
Dati/Oggetti digitali	3	14,3%
Metodologie/Approcci	2	9,5%
Attività	1	4,8%
Istituzioni/Infrastrutture	1	4,8%
Antroponimi	0	0%

Categoria	**N. parole chiave ≥100**	**%**
Progetti	0	0%
Toponimi	0	0%
Totale	**21**	**100%**

Tabella 8. Distribuzione per categoria delle parole chiave con frequenza ≥ 100.

Il quadro risulta in parte diverso rispetto alla distribuzione complessiva del vocabolario. La categoria Concetti/Temi è infatti la più rappresentata (6 parole chiave, 28,6%), seguita da Discipline/Ambiti disciplinari e Tecnologie/Standard/Strumenti (4 parole chiave ciascuna, 19,0%). Al contrario, categorie che nell'intera tassonomia comprendono un numero elevato di parole chiave, come Dati/Oggetti digitali, Metodologie/Approcci e soprattutto Attività, risultano meno rappresentate tra le parole chiave più frequenti.
Questo andamento suggerisce che la numerosità di una categoria non coincide necessariamente con la presenza di termini ad alta ricorrenza, evidenziando come alcuni ambiti siano caratterizzati da un lessico concentrato attorno a pochi concetti chiave, mentre altri presentano una maggiore varietà terminologica.

Estrazione delle co-occorrenze

Per analizzare le relazioni tra i principali temi di ricerca rappresentati nel corpus è stata sviluppata una procedura automatica per il calcolo delle co-occorrenze tra le parole chiave.
L'analisi è stata condotta a partire dai dati grezzi estratti da IRIS e non dal vocabolario normalizzato o dalla tassonomia costruita nelle fasi successive.
Poiché le parole chiave sono registrate nei record nella loro forma originaria, non normalizzata, non è possibile utilizzarle direttamente per il calcolo delle co-occorrenze. È stato quindi predisposto un elenco di corrispondenze tra le forme standard e le relative varianti attestate nel corpus. La procedura utilizza questo elenco per riconoscere le diverse realizzazioni di una stessa parola chiave e ricondurle automaticamente alla forma canonica prima di procedere al conteggio.
L'analisi prende in esame le parole chiave più frequenti del dataset, ovvero quelle con frequenza >100 (Figura 9). Poiché il conteggio è calcolato a partire dalle pubblicazioni e non dalla tassonomia, il numero di parole chiave selezionate (13) non coincide con quello ottenuto nelle analisi precedenti.
Per ciascuna di esse vengono identificate tutte le pubblicazioni in cui compare la forma canonica e conteggiate le co-occorrenze con tutte le altre parole chiave presenti nello stesso record. La procedura restituisce quindi, per ogni parola chiave, la distribuzione completa delle co-occorrenze con l'intero vocabolario del corpus, calcolando sia il numero di pubblicazioni condivise, sia la frequenza relativa della co-occorrenza, espressa come percentuale rispetto al totale delle pubblicazioni contenenti la parola chiave selezionata.

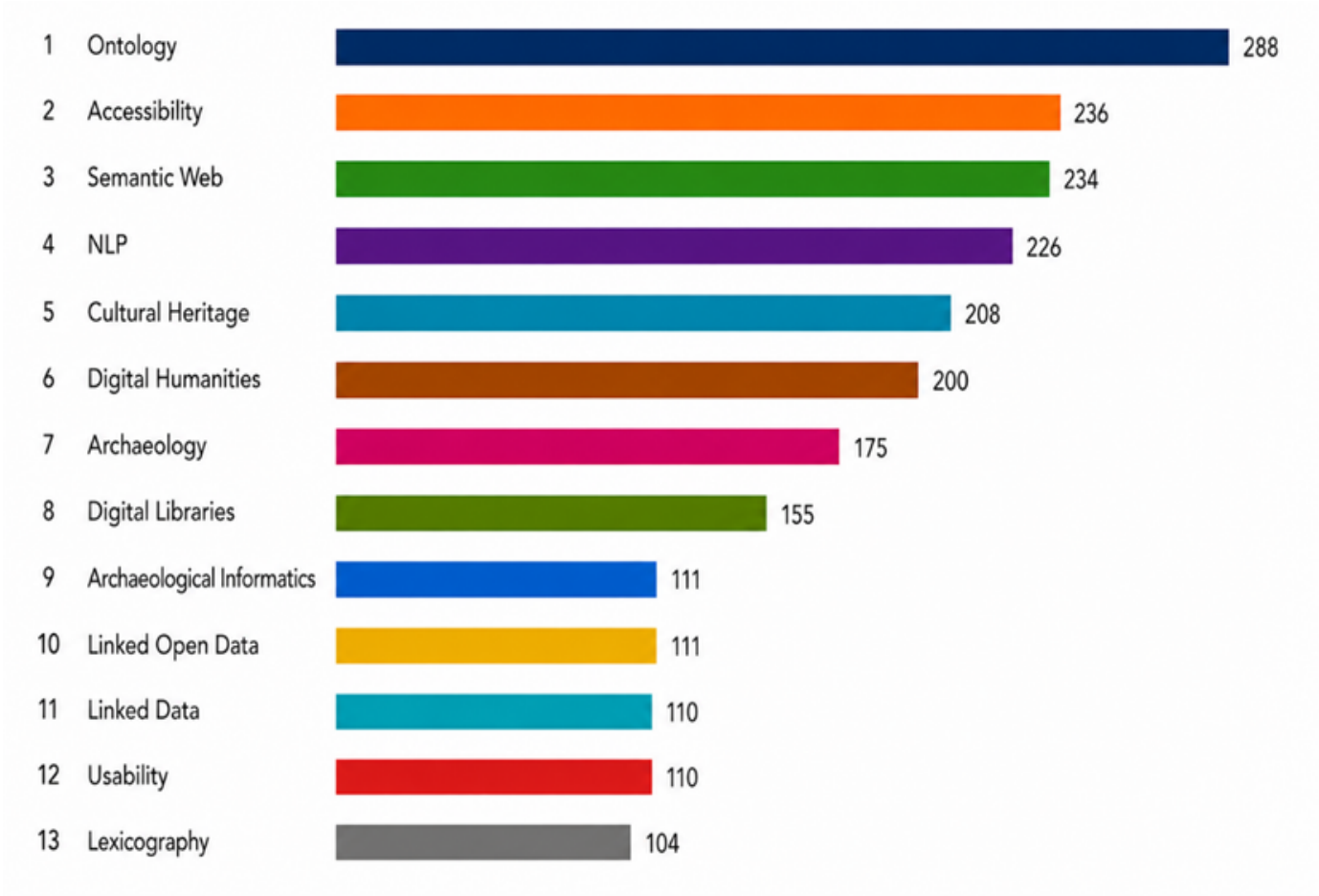

Figura 9. Parole chiave con frequenza > 100 nel dataset delle pubblicazioni.

Prima di analizzare le relazioni di co-occorrenza, è opportuno soffermarsi brevemente sulle varianti associate alle parole chiave selezionate.

La loro ricognizione consente infatti di osservare le strategie denominative adottate dalla comunità scientifica, evidenziando ad esempio la preferenza tra forme inglesi o italiane, e l'impiego degli acronimi. In particolare, si osserva una netta prevalenza dei prestiti inglesi rispetto ai corrispondenti traducenti italiani (*Digital Humanities* 191 occ. - *Informatica umanistica* 8 occ.; *Cultural Heritage* 170 occ. - *patrimonio culturale* 10 occ.; *Ontology* 161 occ. - *Ontologia* 5 occ., ecc.), confermando il predominio della terminologia internazionale nel dominio delle DH.

Per quanto riguarda gli acronimi, si nota una preferenza per la formulazione estesa rispetto alla sigla: *NLP* risulta meno frequente (38 occ.) rispetto a *Natural Language Processing* (142), che è prevalente, ma più attestato rispetto alla forma italiana *Trattamento Automatico del Linguaggio* (12), indice di una terminologia ormai stabilizzata ma ancora espressa attraverso convenzioni differenti. Parimenti, l'acronimo LOD occorre soltanto 3 volte, mentre *Linked Open Data* presenta una più ampia circolazione (100 occ.).

Archaeology rappresenta un'eccezione rispetto alla tendenza generale: è infatti l'unico caso in cui il traducente italiano prevale sulla forma inglese (133 occ. contro 33 occ.).

Le principali co-occorrenze delle parole chiave a maggiore frequenza sono presentate nella Tabella 9.

Parole chiave	**Principale co-occorrenza**	**N. co-occorrenze**	**%**
Ontology	Semantic Web	102	35,42%
Accessibility	Usability	96	40,68%
Semantic Web	Ontology	102	43,59%
NLP	Ontology	24	10,62%
Cultural Heritage	Linked Open Data	31	14,90%
Digital Humanities	Semantic Web	24	12,00%
Archaeology	Archaeological Informatics	107	61,14%
Digital Libraries	Semantic Web	36	23,23%
Archaeological Informatics	Archaeology	107	96,40%
Linked Open Data	Cultural Heritage	31	27,93%
Linked Data	Semantic Web	50	45,45%
Usability	Accessibility	96	87,27%
Lexicography	Lessico	31	29,81%

Tabella 9. Principali co-occorrenze delle parole chiave a maggiore frequenza.

Le co-occorrenze evidenziano diverse tipologie di relazione tra le parole chiave. In alcuni casi la co-occorrenza riflette l'appartenenza allo stesso dominio disciplinare, come per *Archaeological Informatics* - *Archaeology* (107 pubblicazioni condivise). In altri, la correlazione è legata al fatto che i due termini sono riferibili allo stesso campo concettuale: è il caso di *Ontology* e *Semantic Web*, oppure di *Accessibility* e *Usability*.
Per visualizzare in modo più immediato tali relazioni, la Figura 10 riporta, per ciascuna delle parole chiave di riferimento, le cinque co-occorrenze più frequenti.

Figura 10. Principali co-occorrenze per ciascuna parola chiave.

Per valutare l'ampiezza delle relazioni delle parole chiave più frequenti con il resto del vocabolario, è stato calcolato il numero di parole distinte con cui ciascuna di esse co-occorre. Sono state considerate soltanto quelle con frequenza ≥ 3, così da escludere associazioni occasionali (Tabella 10).

Parole chiave	**N. occ.**	**Parole chiave distinte co-occorrenti**	**Parole chiave di riferimento co-occorrenti (≥ 3)**
Ontology	288	366	7
Accessibility	236	231	1
Semantic Web	234	255	8
NLP	226	321	5
Cultural Heritage	208	358	7
Digital Humanities	200	241	9
Archaeology	175	168	3
Digital Libraries	155	148	5

Parole chiave	N. occ.	Parole chiave distinte co-occorrenti	Parole chiave di riferimento co-occorrenti (≥ 3)
Archaeological Informatics	111	22	2
Linked Open Data	111	124	6
Linked Data	110	137	4
Usability	110	105	1
Lexicography	104	98	4

Tabella 10. Ampiezza delle relazioni delle parole chiave a maggiore frequenza.

Tra le parole chiave più frequenti del corpus, *Ontology* e *Digital Humanities* meritano un approfondimento specifico per il diverso ruolo che assumono nella rete delle co-occorrenze. La prima per la sua rilevanza quantitativa: *Ontology* risulta infatti il termine più frequente del dataset e presenta il più ampio insieme di co-occorrenze. *Digital Humanities*, oltre a rappresentare il dominio di riferimento della tassonomia, si distingue invece per l'elevato numero di relazioni con le altre parole chiave di riferimento, configurandosi come un termine trasversale ai principali ambiti della ricerca.
Ontology presenta una marcata associazione con *Semantic Web* (102 co-occorrenze, pari al 35,4% delle occorrenze della parola chiave). Le successive co-occorrenze risultano nettamente meno frequenti (dal 10,8% al 6,6%).
Diversamente da *Ontology*, *Digital Humanities* non presenta una singola co-occorrenza predominante, ma associazioni significative con numerose parole chiave. Le cinque co-occorrenze più frequenti mostrano infatti valori relativamente vicini tra loro (da 24 a 13 co-occorrenze; 12% - 6,5%).

3. Questionario sulle Digital Humanities al CNR DSU

L'obiettivo del questionario è stato quello di ampliare i dati ottenuti con la precedente analisi (Sezione 1). Avendo infatti centrato quest'ultima sul reperimento di dati tramite piattaforme bibliografiche, si è pensato utile integrare i risultati facendo appello diretto ai colleghi e alle colleghe del DSU.

Si è trattato di un questionario piuttosto breve, con un tempo di compilazione stimato intorno ai 5 minuti, in cui si è chiesto ai partecipanti di segnalare (oltre al riferimento al proprio nome/cognome, istituto CNR di appartenenza e ORCID):

1. le **aree** di afferenza disciplinare, secondo sia la più recente classificazione dell'*European Research Council* (ERC), sia le tabelle dei settori scientifico-disciplinari del Ministero dell'Università e della Ricerca (MUR). In particolare, è stato chiesto di indicare la propria

area rispetto alle attività di ricerca in ambito DH al di là dell'area ERC *SH5_11 Digital humanities; digital approaches to literary studies and philosophy*, considerando quest'ultima come area di riferimento della comunità;
2. un massimo di 3 **keyword** per identificare la propria tematica di ricerca;
3. un massimo di 3 **eventi** DH (conferenze, workshop) frequentati negli ultimi tre anni (2024–2026 in corso);
4. un massimo di 3 **riviste** in cui sono stati pubblicati i propri contributi DH;
5. un massimo di 3 **pubblicazioni** riconducibili a tematiche di DH.

Il questionario è stato realizzato tramite *form* di Microsoft 365 ed è stato disponibile dal 1° al 20 luglio 2026. Complessivamente sono state raccolte 49 risposte. Si riportano di seguito solo alcuni dei dati raccolti come campione preliminare per successive analisi.

Persone e Istituti. I partecipanti provengono dai seguenti Istituti del DSU, confermando che le DH rappresentano una competenza trasversale all'interno del Dipartimento:
- IGSG - Istituto di Informatica Giuridica e Sistemi Giudiziari;
- ILC - Istituto di Linguistica Computazionale "Antonio Zampolli";
- IRFiS - Istituto per la Ricerca Filosofica e Storica;
- IRPPS - Istituto di Ricerche sulla Popolazione e le Politiche Sociali;
- ISEM - Istituto di Storia dell'Europa Mediterranea;
- ISMed - Istituto di Studi sul Mediterraneo;
- ISPC - Istituto di Scienze del Patrimonio Culturale;
- ISTC - Istituto di Scienze e Tecnologie della Cognizione;
- ITD - Istituto per le Tecnologie Didattiche;
- OVI - Opera del Vocabolario Italiano.

L'ISPC rappresenta il nucleo più numeroso del campione con 14 partecipanti al questionario, seguito dall'ILC con 9 partecipanti, ITD con 6, ISTC con 5, IRFiS con 4, ISMed, IGSG e IRPPS con 3 ciascuno, OVI e ISEM con un partecipante.
Incrociando gli ORCID raccolti tramite il questionario con i dati analizzati nelle sezioni precedenti, sono 21 i partecipanti al questionario i cui dati rientrano nella "Lista AIUCD" (autori o co-autori di contributi presentati alle edizioni della conferenza AIUCD negli ultimi dieci anni, 2016–2025); di questi, 16 sono anche presenti nella "Lista SH5_11" (ricercatori CNR che nell'ultima mappatura istituzionale dell'afferenza alle aree disciplinari hanno dichiarato fra i propri settori di interesse il settore SH5_11). Dei rimanenti 28 partecipanti, 2 non hanno inserito il proprio ORCID e 26 rispondenti con ORCID non sono presenti nella lista degli autori AIUCD.

Keyword. Il questionario censisce circa 100 keyword, di cui una sessantina sono quelle inserite dai 26 ricercatori non censiti nella precedente analisi. I risultati sono complessivamente allineati

con quanto discusso precedentemente (Sezione 2), rappresentando una sottoparte dei dati già presentati.

Guardando nel dettaglio le keyword inserite dai 26 ricercatori non censiti dalla precedente analisi, un primo asse di ricerca particolarmente evidente riguarda il patrimonio culturale (digitale), con temi quali *Digital Heritage*, *Digital Archaeology*, *Digital Critical Edition*, *Virtual Architecture*, *History of Illuminated Manuscripts*, *Paleography* e altro ancora. A questo filone si affianca una forte attenzione verso le tecnologie per la documentazione, l'analisi e la valorizzazione del patrimonio, come testimoniano keyword quali *3D Reconstruction*, *3D Analysis*, *Archaeological GIS*, *Heritage Building Information Modelling*, *Deep Mapping*, *Immersive Technology*, *Virtual Reality* e *Mixed Reality*. Un secondo asse è rappresentato dall'intelligenza artificiale con keyword come *Human-Centred Artificial Intelligence*, *LLM*, *Knowledge Representation and Reasoning*, *Ontologies*, *Ontology-based Data Access*, *Semantic Web* e *Semantic Web for Cultural Heritage*.

Anche in questo campione è dunque presente un forte interesse per la costruzione di modelli formali della conoscenza e di sistemi capaci di garantire interoperabilità, interpretabilità e riuso dei dati. Un ulteriore filone riguarda le applicazioni dell'intelligenza artificiale e delle tecnologie digitali ai domini della **didattica**, del **diritto** e delle **scienze sociali**, come mostrano le keyword *Educational Technology*, *Digital Holocaust Education*, *Legal Knowledge Representation*, *Policy Intelligence*, *Social Computing* e *Social Politics and Movements*. Infine, emergono temi trasversali quali *Open Data*, *Heritage Data Integration*, *Human-Technology Interaction*, *User-centered Design*, *Gender Studies*, *Environmental Digital Humanities* e *Digital Ecology*, che testimoniano una crescente attenzione agli aspetti sociali, etici e partecipativi nelle ricerche che utilizzano metodologie e tecnologie digitali.

Pur trattandosi di considerazioni preliminari, emerge una marcata convergenza tra ricerca sul patrimonio culturale e tecnologie dell'intelligenza artificiale, segno che la gestione, l'integrazione e l'analisi dei dati culturali continuano a rappresentare uno dei principali terreni di ricerca in ambito DH. Inoltre, la presenza di temi legati al diritto, all'educazione, alla partecipazione sociale e alle questioni ambientali suggerisce che le DH, almeno nel campione considerato, stanno progressivamente ampliando il proprio raggio d'azione verso problemi di rilevanza sociale, confermando la loro natura di ambito di ricerca intrinsecamente interdisciplinare e orientato all'impatto.

Aree ERC. La Figura 11 mostra le aree di afferenza ERC dei 26 rispondenti al questionario non presenti nella precedente analisi.

Partecipazione alla comunità scientifica. L'analisi degli eventi frequentati nel triennio 2024–2026 (in corso) offre un'indicazione significativa delle principali comunità scientifiche di riferimento.

Il convegno annuale dell'*Associazione per l'Informatica Umanistica e la Cultura Digitale* (AIUCD) rappresenta il principale punto di aggregazione nazionale, confermandosi come il

luogo di incontro fondamentale della comunità italiana delle DH. Accanto ad AIUCD emergono diversi eventi nazionali e internazionali, tra cui *Digital Heritage* e la conferenza *DH* dell'*Alliance of Digital Humanities Organization* (ADHO), così come numerosi seminari, workshop e conferenze dedicate a specifiche tematiche d'interesse per i singoli ricercatori e ricercatrici. I dati raccolti mostrano inoltre un'attiva partecipazione del CNR agli eventi legati alle infrastrutture europee quali, ad esempio, CLARIN e DARIAH.

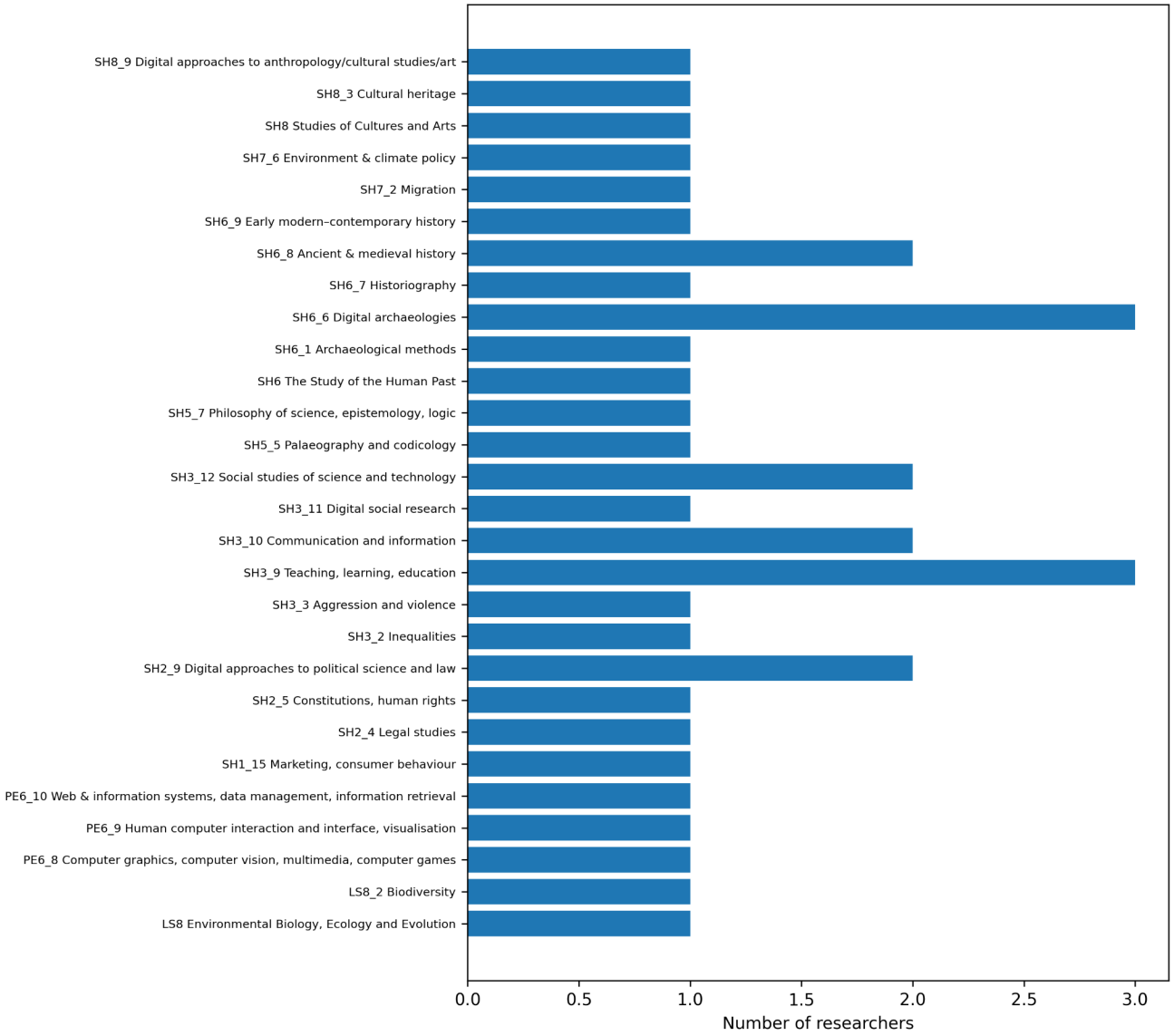

Figura 11. Aree ERC dei 26 partecipanti al questionario.

Riviste. La rivista *Umanistica Digitale,* coordinata da AIUCD, risulta la sede editoriale maggiormente rappresentata, seguita da *Archeologia e Calcolatori.* Seguono poi tante altre riviste a carattere sia nazionale sia internazionale. Nel complesso l'elenco (Figura 12) comprende riviste specializzate in rappresentazione della conoscenza, patrimonio culturale,

diritto dell'intelligenza artificiale, linguistica, archeologia, scienze dell'informazione, scienze della formazione, studi sulla memoria sociale e altro ancora.

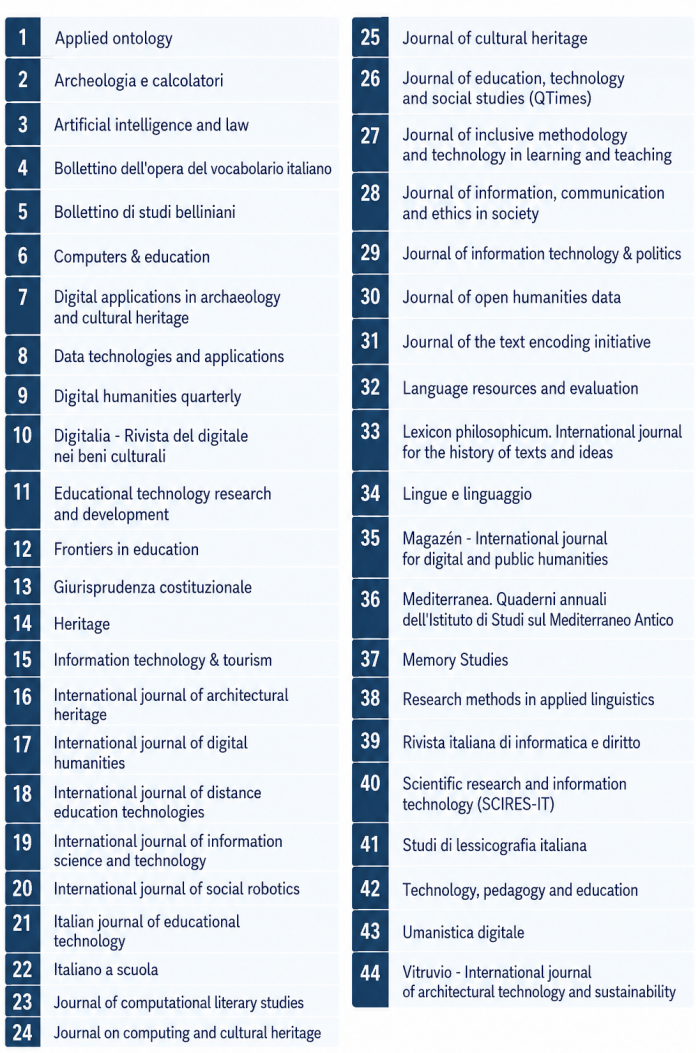

Figura 12. Elenco di riviste emerso dal questionario.

Conclusioni

L'indagine condotta si colloca ancora a un livello di analisi preliminare; i dati raccolti richiedono infatti un ulteriore lavoro di elaborazione per ottenere risultati più robusti, sia sul piano tecnico (ad esempio attraverso processi di normalizzazione), sia su quello dell'analisi quantitativa e qualitativa. Ciò nonostante, l'indagine restituisce il profilo di una comunità scientifica del CNR estremamente articolata.

In particolare, le DH non appaiono concentrate in un unico settore disciplinare né in un singolo istituto. Al contrario, costituiscono una rete trasversale che coinvolge competenze linguistiche, filologiche, giuridiche, storico-archeologiche, informatiche, filosofiche, documentarie e pedagogiche, solo per citarne alcune. L'elevata partecipazione agli eventi AIUCD testimonia l'esistenza di una solida comunità nazionale di riferimento, mentre la presenza di numerose

conferenze e riviste internazionali evidenzia una marcata apertura al panorama internazionale della ricerca.

Nel loro insieme, i dati delineano una comunità dinamica e pienamente inserita nelle principali reti di ricerca nazionali e internazionali, confermando il ruolo delle DH come ambito strategico di innovazione metodologica e di collaborazione tra discipline.

Proprio in ragione del carattere preliminare dello studio, rimane ancora molto lavoro da svolgere. In particolare, è nostro obiettivo pubblicare i dati raccolti, sia nel loro formato "grezzo" sia in una versione normalizzata, affinché possano essere consultati e riutilizzati dalla comunità scientifica per ulteriori studi e analisi.

Fonti dei dati

Associazione per l'Informatica Umanistica e la Cultura Digitale (AIUCD), Dipartimento di Studi Linguistici e Culturali Comparati, Università Ca' Foscari, Ca' Bembo – Dorsoduro 1075, Fondamenta Toffetti, 30123, Venezia. https://www.aiucd.it.

Consiglio Nazionale delle Ricerche (CNR). Piattaforma IRIS (Institutional Research Information System). https://iris.cnr.it/.

Un ponte FAIR per le Digital Humanities: l'integrazione tra OpenAIRE e DH-ATLAS al CNR

A FAIR Gateway for Digital Humanities: OpenAIRE and DH-ATLAS Integration at CNR

Angelo Mario Del Grosso, CNR-ILC, angelomario.delgrosso@cnr.it

Abstract
Le Digital Humanities richiedono strumenti capaci di collegare persone, prodotti, progetti e temi oltre i confini disciplinari. Il contributo mostra come i progetti DH-ATLAS e OpenAIRE integrino granularità disciplinare e interoperabilità per rendere la ricerca DH reperibile, interrogabile e riusabile.

Abstract (English)
Digital Humanities require tools that connect people, outputs, projects, and topics across disciplinary boundaries. This brief shows how DH-ATLAS and OpenAIRE projects combine disciplinary granularity and interoperability to make DH research more findable, queryable and reusable.

Keywords
digital humanities – knowledge graph – interoperability – open science – cultural heritage

Introduzione

Nel DSU del CNR operano oggi 13 Istituti[1], con competenze che spaziano dalla linguistica computazionale alla filologia e al patrimonio culturale. La configurazione del Digital Humanities Center (DHC)[2], nato per aggregare, promuovere e valorizzare le attività DH dell'Ente e creare una rete visibile e dinamica, conferma che questa comunità non coincide con una singola disciplina né con un singolo istituto, ma emerge dall'interazione tra metodi d'indagine, oggetti di ricerca e competenze scientifiche differenti.

Ricostruire l'ecosistema della ricerca DH significa rappresentare le relazioni semanticamente significative tra persone, prodotti, progetti, istituzioni, metodi, servizi e oggetti culturali. Una prima ricognizione basata prevalentemente sui dati della piattaforma CNR-IRIS[3], approfondita nel contributo precedente (Agnoloni, Francalanci, Sanfilippo), mostra che affiliazioni, metadati bibliografici e parole chiave restituiscono soltanto una rappresentazione parziale della comunità

[1] https://www.cnr.it/it/dipartimento/514/istituti/scienze-umane-e-sociali-patrimonio-culturale (ultima consultazione dei siti web: 5/10/2026).

[2] https://www.dsu.cnr.it/digital-humanities-center/

[3] https://iris.cnr.it/

DH presente al CNR. Ne deriva l'esigenza metodologica di integrare le fonti disponibili in un modello capace di rendere esplicita la rete di relazioni che caratterizza la ricerca DH.

Dalla modellazione semantica all'interoperabilità

Per rispondere a questa esigenza occorre rendere espliciti il significato delle entità e le relazioni che le connettono. Ontologie e Knowledge Graph (KG) consentono di integrare dati eterogenei in una rappresentazione esplicita e condivisa della conoscenza, favorendo in particolare l'interoperabilità e il riuso dei dati secondo i principi FAIR (*Findable, Accessible, Interoperable, Reusable*)[4]. In *Knowledge Representation* un'ontologia formalizza concetti, relazioni e vincoli di un dominio (Guarino et al. 2009), mentre un KG organizza entità e relazioni in una struttura interrogabile (Hogan et al. 2021), eventualmente governata da un modello ontologico di riferimento. Un prodotto DH può così essere collegato ad autori, istituzioni, progetti, metodi, obiettivi e servizi.

Il progetto DH-ATLAS[5] nasce dai limiti dei modelli generalisti nel rappresentare con sufficiente granularità gli output delle Digital Humanities (Giacomini et al. 2026). L'ontologia ATLAS riusa classi e proprietà di modelli e vocabolari ampiamente adottati: Schema.org[6] costituisce il *backbone*, integrato da DCTerms[7] e FaBiO[8] per aumentare la granularità descrittiva e bibliografica. Per i prodotti della ricerca ATLAS distingue tipologie specifiche quali *Digital Scholarly Edition*, *Text Collection*, *Linked Open Data*, *Ontology*, *Software*, *Language Model* e *3D Digital Twin* (Bardi et al. 2026). La piattaforma ATLAS è stata realizzata estendendo CLEF (*Crowdsourcing Linked Entities via web Form*)[9], con funzionalità di *data entry*, riconciliazione ed estrazione di entità. Il KG consente così di collegare prodotti, persone, organizzazioni, progetti, siti e servizi attraverso identificatori e vocabolari condivisi. Il grafo può essere esplorato per tipologia di risorsa, persone, organizzazioni e progetti e interrogato tramite API (Figure 1a e 1b).

[4] https://www.go-fair.org/fair-principles/

[5] https://dh-atlas.github.io/project.html

[6] https://schema.org/

[7] https://www.dublincore.org/specifications/dublin-core/

[8] https://www.sparontologies.net/ontologies/fabio

[9] https://github.com/polifonia-project/clef

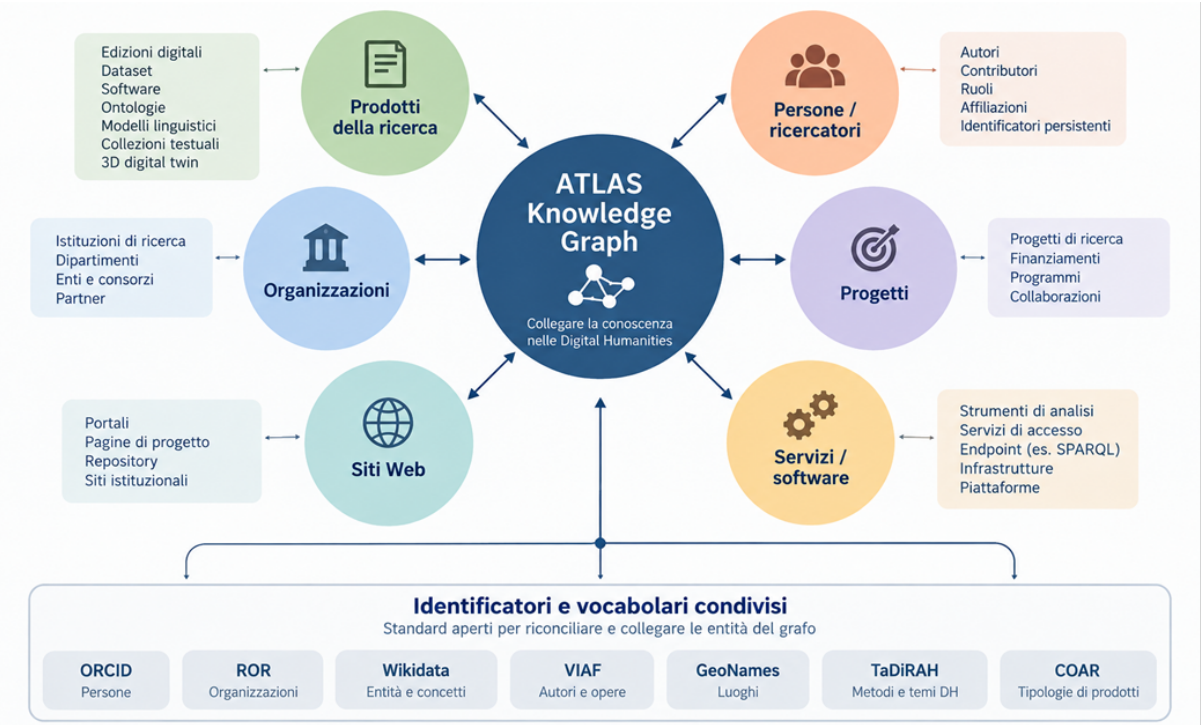

Figura 1a. DH-ATLAS Knowledge Graph.

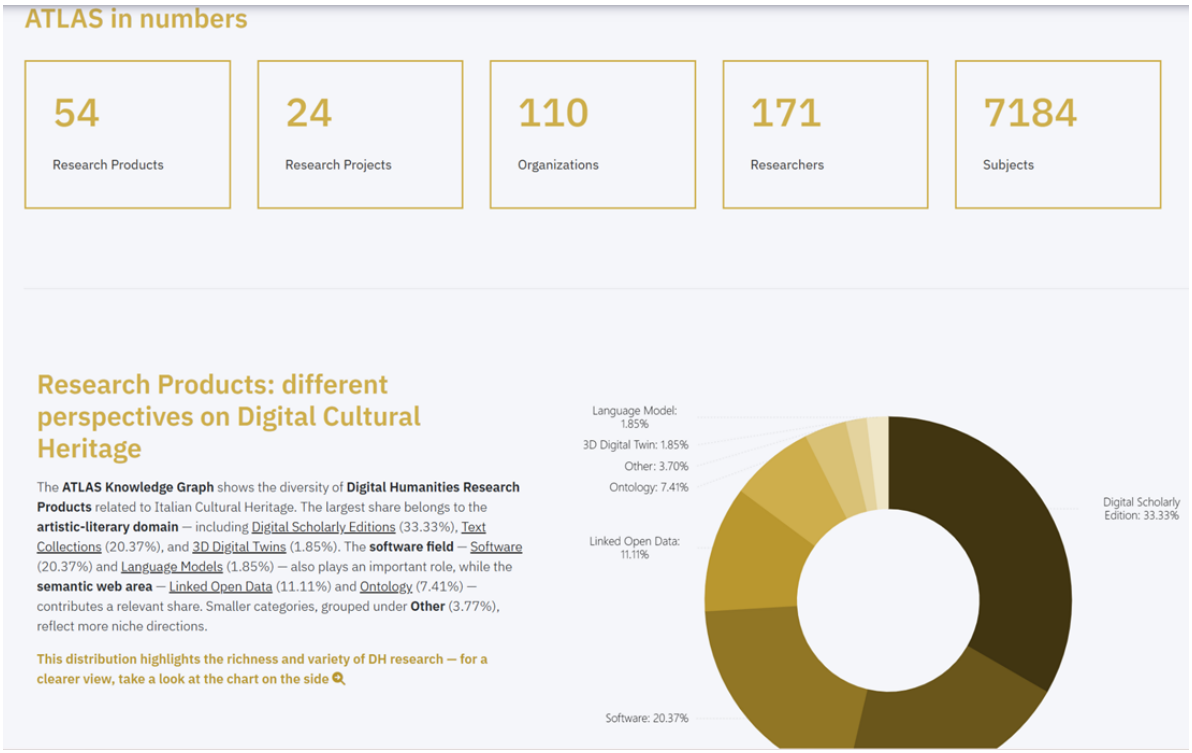

Figura 1b. Dashboard ATLAS: consistenza del Knowledge Graph e distribuzione dei research products per tipologia.

OpenAIRE, a sua volta, opera a una scala complementare[10]. L'OpenAIRE Graph è uno *scholarly knowledge graph* aperto che collega prodotti della ricerca, persone, organizzazioni, progetti, finanziamenti, fonti dati e comunità. Tra i vari servizi offerti, CONNECT[11] consente alle comunità scientifiche di creare *Community Gateways*[12], cioè portali tematici configurabili costruiti sullo *Scholarly KG*. Tra questi il *Gateway Digital Humanities and Cultural Heritage*[13] è dedicato alle Digital Humanities e al Cultural Heritage e comprende anche la comunità italiana DH-CH (Figura 2).

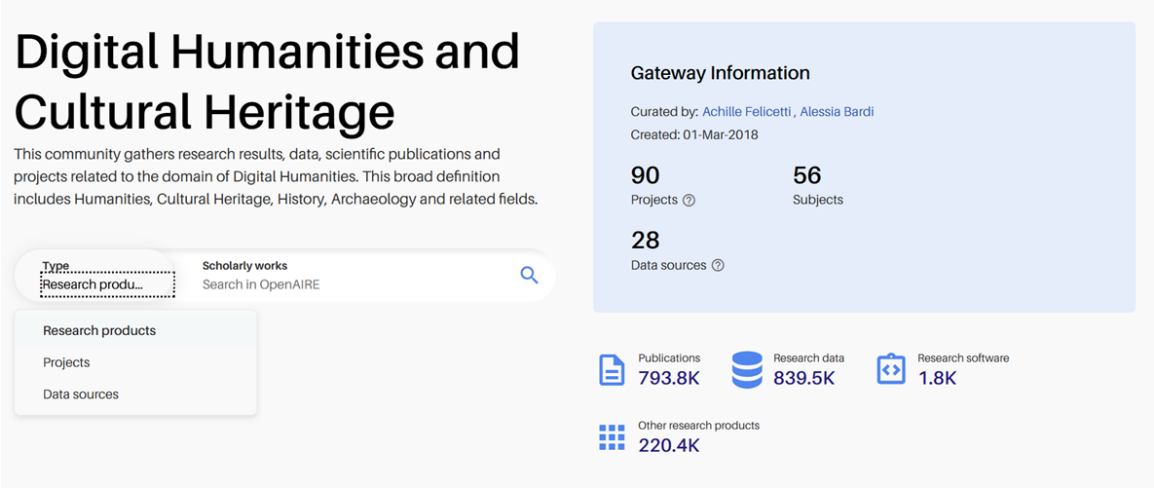

Figura 2. Gateway Digital Humanities and Cultural Heritage (DH-CH) di OpenAIRE CONNECT.

I due livelli sono dunque complementari, ma non sovrapponibili. Per integrare ATLAS in OpenAIRE, un nucleo trasversale dei metadati viene mappato sulle OpenAIRE Guidelines ed esposto tramite OAI-PMH[14] per la raccolta automatica. I tipi di prodotto ATLAS sono inoltre allineati ai COAR Resource Types[15], adottati da OpenAIRE come vocabolario controllato. Per esempio, *Digital Scholarly Edition* diventa *encodedData* e *Ontology* diventa *KnowledgeOrganizationSystem*. Un servizio *RESTful* configurato con RAMOSE[16] e distribuito nell'infrastruttura D4Science[17] affianca il processo di integrazione e interrogazione favorendo il relativo endpoint[18] ad esporre operazioni predefinite che consentano di accedere al grafo senza formulare direttamente query SPARQL (Rubin et al. 2026). L'interoperabilità agisce quindi come mediazione controllata tra ricchezza semantica locale e infrastrutture comuni.

[10] https://www.openaire.eu/
[11] https://connect.openaire.eu/
[12] https://www.openaire.eu/research-community-gateway-guide
[13] https://dh-ch.openaire.eu/
[14] https://www.openarchives.org/pmh/
[15] https://vocabularies.coar-repositories.org/
[16] https://opencitations.github.io/ramose/intro.html
[17] https://www.d4science.org/
[18] https://atlas-prin2022-api.cloud.d4science.org/

Due casi d'uso: comunità e risorse

I due casi d'uso seguenti mostrano come la combinazione di modellazione semantica e interoperabilità possa sostenere le attività di mappatura e valorizzazione promosse dal DHC. Il primo riguarda la mappatura delle DH al CNR. In un ambiente di rappresentazione della conoscenza basato su KG, persone, affiliazioni, prodotti e temi possono essere collegati lungo percorsi del tipo *ricercatore > istituto > prodotto > tema/metodo > progetto/servizio*. La domanda di ricerca può quindi essere formulata cercando "quali ricercatori CNR lavorano sulle ontologie applicate al Cultural Heritage, in quali strutture e con quali prodotti". Un'interrogazione costruita su tali relazioni non si limita alla presenza di una parola chiave, ma incrocia entità connesse e consente di impostare una mappatura più affidabile e aggiornabile delle attività DH dell'Ente.

Il secondo caso d'uso riguarda il progetto *Bellini Digital Correspondence* (BDC)[19], l'edizione scientifica digitale delle lettere autografe di Vincenzo Bellini conservate al Museo civico Belliniano di Catania. L'edizione, codificata secondo le linee guida TEI e resa fruibile via Web, comprende 40 unità testuali corrispondenti a 35 unità codicologiche e 111 immagini digitali e offre indici di persone, luoghi, enti, opere e terminologia musicale (Del Grosso and Spampinato 2023). In ATLAS, BDC è descritta come *Digital Scholarly Edition* e può essere collegata al progetto, ai ricercatori, alle organizzazioni e agli strumenti che ne permettono la fruizione (Figura 3). Nel Gateway DH-CH, una ricerca per il titolo "Bellini Digital Correspondence" restituisce inoltre prodotti scientifici pertinenti, ampliando il contesto di ricerca oltre il singolo record (Figura 4).

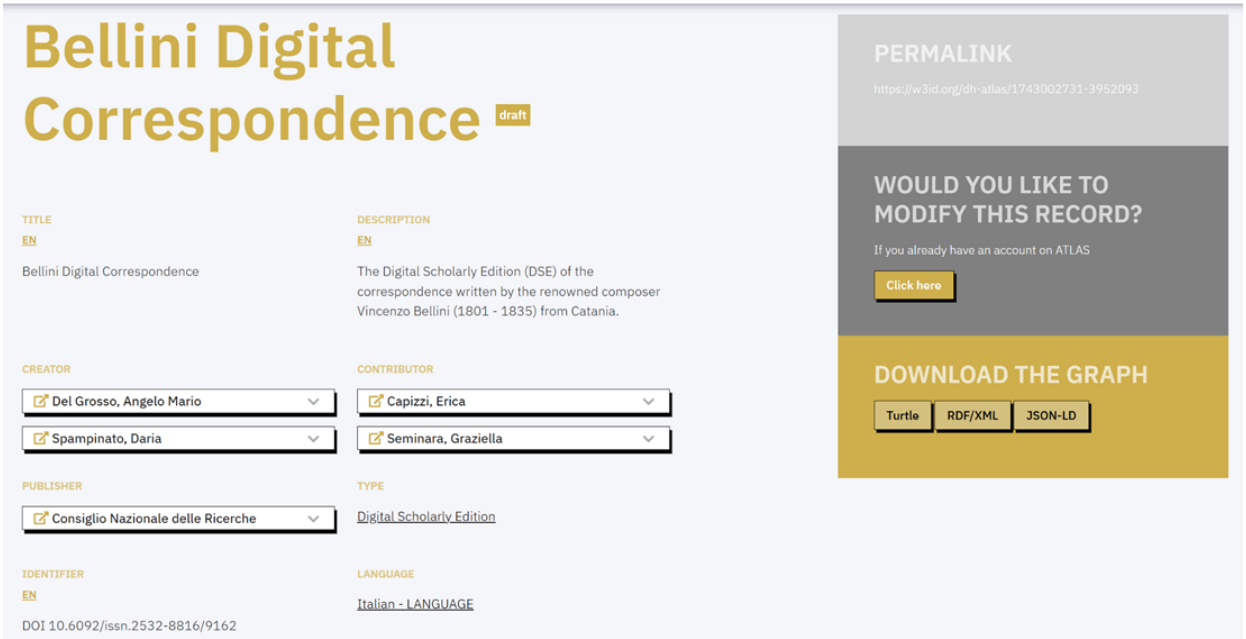

Figura 3. *Bellini Digital Correspondence* in ATLAS: tipologia del prodotto e collegamenti a creatori e organizzazioni.

[19] https://bellinicorrespondence.cnr.it/

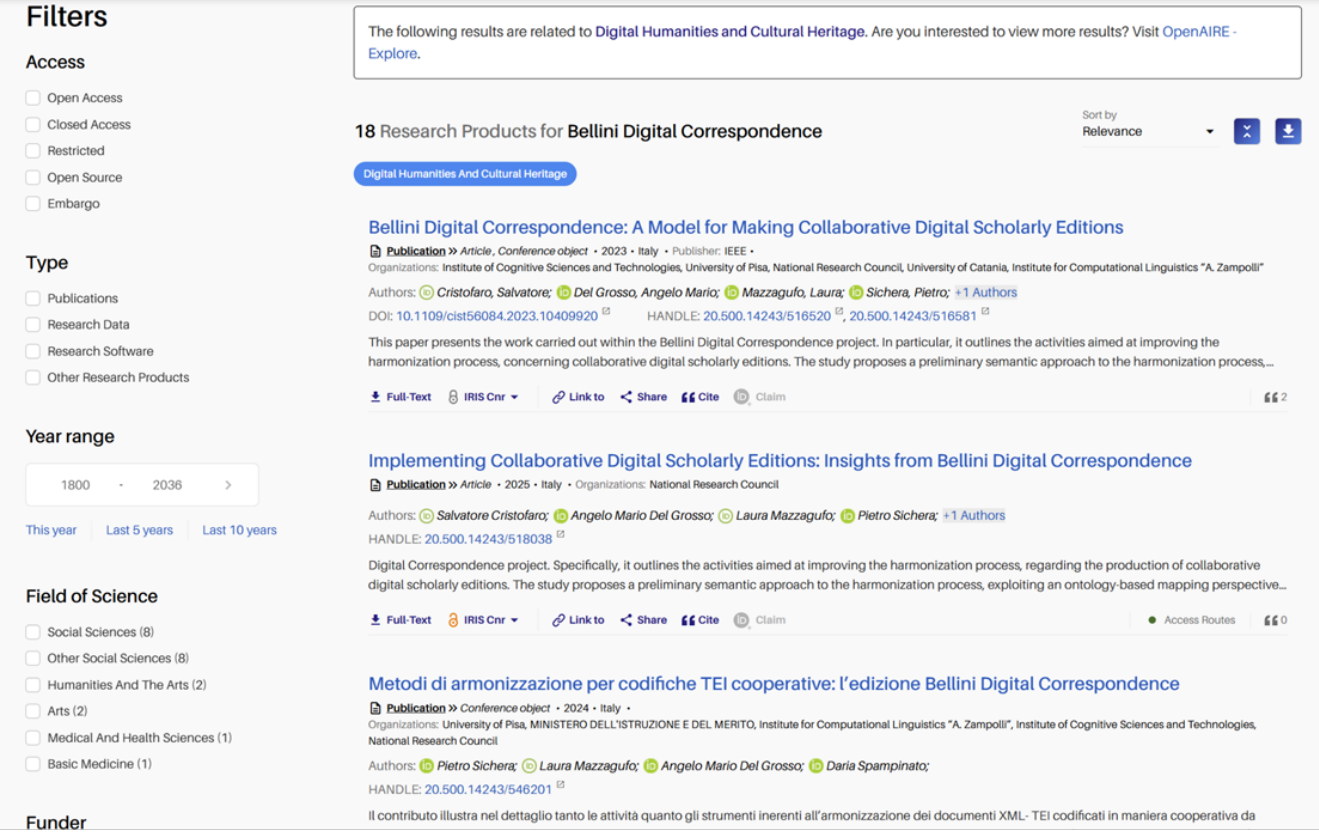

Figura 4. Ricerca dei prodotti relativi a "Bellini Digital Correspondence" nel Gateway DH-CH di OpenAIRE.

La stessa logica guida la sottocomunità italiana DH-CH. Le informazioni estratte dal KG di ATLAS sono state utilizzate per definire criteri computabili con cui riconoscere prodotti pertinenti al patrimonio culturale italiano (Rubin et al. 2026).
In CONNECT tali criteri possono combinare, per esempio, soggetti e parole chiave, affiliazioni e finanziamenti. Un grafo specialistico può così fornire conoscenza di dominio a un'infrastruttura più ampia, migliorando la scoperta di prodotti e relazioni senza appiattire la specificità disciplinare.

Conclusioni

Per il CNR, mappare in modo affidabile l'ambito delle Digital Humanities significa ricostruire una rete di relazioni tra persone, temi, progetti, metodi e servizi, rappresentata attraverso dati interoperabili e conformi ai principi FAIR.
ATLAS offre granularità disciplinare, mentre OpenAIRE offre scala, aggregazione e reperibilità trasversale. La loro integrazione mostra come si possa passare dalla catalogazione dei risultati alla rappresentazione delle relazioni che definiscono le comunità che li producono.

In questa prospettiva, il DHC può agire come punto di raccordo tra dati istituzionali, modelli semantici e infrastrutture aperte, contribuendo a descrivere, interrogare e connettere la comunità DH del CNR[20].

Riferimenti bibliografici

Bardi, Alessia, Marina Buzzoni, Marilena Daquino, Riccardo Del Gratta, Angelo Mario Del Grosso, Roberto Rosselli Del Turco, Franz Fischer, et al. 2026. *The ATLAS Whitebook and Guidelines*. Quaderni di Umanistica Digitale. Associazione per l'Informatica Umanistica e la Cultura Digitale – AIUCD. https://doi.org/10.6092/unibo/amsacta/8808.

Giacomini, Sebastiano, Chiara Martignano, Giorgia Rubin, Alessia Bardi, Marina Buzzoni, Marilena Daquino, Riccardo Del Gratta, et al. 2026. "ATLAS: A Knowledge Graph to Enhance the Findability of International Scholarly Research on the Italian Digital Cultural Heritage." *Umanistica Digitale* 10 (22): 45–76. https://doi.org/10.60923/issn.2532-8816/21765.

Del Grosso, Angelo Mario, and Daria Spampinato. 2023. *Bellini Digital Correspondence*. CNR Edizioni. https://bellinicorrespondence.cnr.it.

Guarino, Nicola, Daniel Oberle, and Steffen Staab. 2009. "What Is an Ontology?" In *Handbook on Ontologies*, edited by Steffen Staab and Rudi Studer. Springer. https://doi.org/10.1007/978-3-540-92673-3_0.

Hogan, Aidan, Eva Blomqvist, Michael Cochez, Claudia d'Amato, Gerard de Melo, Claudio Gutierrez, Sabrina Kirrane, et al. 2021. "Knowledge Graphs." *ACM Computing Surveys* 54 (4): 1–37. https://doi.org/10.1145/3447772.

Rubin, Giorgia, Alessia Bardi, Riccardo Del Gratta, and Angelo Mario Del Grosso. 2026. "Enhancing Interoperability of SPARQL Endpoints: RESTful and OAI-PMH API for the DH-ATLAS Project." In *Proceedings of the 22nd Conference on Information and Research Science Connecting to Digital and Library Science (IRCDL 2026)*. CEUR Workshop Proceedings 4240. https://ceur-ws.org/Vol-4240/paper38.pdf.

[20] Dichiarazione sull'uso dell'intelligenza artificiale: il testo è stato sottoposto a revisione linguistica con il supporto di strumenti di intelligenza artificiale generativa. La responsabilità dei contenuti resta dell'autore.

Nodi, reti e sinergie

Nodes, Networks, and Synergies

Cristina Marras, CNR-IRFiS, cristina.marras@cnr.it
Alessia Scognamiglio, CNR-IRFiS, alessia.scognamiglio@cnr.it

Abstract
A partire dalla prima mappatura delle Digital Humanities al CNR, con particolare attenzione alle attività di ricerca del DSU, presentata in questo numero dei Brief il presente contributo propone il Centro di Digital Humanities del CNR-DSU come nodo strategico per valorizzare il capitale umano, rafforzare le reti internazionali, promuovere interoperabilità, scienza aperta, public engagement e buone pratiche metodologiche, e sviluppare nuove sinergie tra Digital Humanities e Intelligenza Artificiale. Il contributo delinea così un framework strategico volto a consolidare il ruolo del CNR nell'ecosistema globale delle Digital Humanities.

Abstract (English)
Building on the first mapping of Digital Humanities at the CNR, with particular attention to the research activities of the DSU, presented in this issue of the *Briefs*, this contribution proposes the CNR-DSU Digital Humanities Centre as a strategic hub for enhancing human capital, strengthening international networks, promoting interoperability, Open Science, public engagement, and methodological best practices, and developing new synergies between Digital Humanities and Artificial Intelligence. The contribution thus outlines a strategic framework aimed at consolidating the CNR's role within the global Digital Humanities ecosystem.

Keywords
digital humanities – networks – synergies – strategies

1. Nodi

Dall'accurato lavoro di mappatura e dall'analisi sistematica dei dati presentati nel contributo *Mappare le Digital Humanities al CNR DSU attraverso i dati della ricerca: pubblicazioni, autori e parole chiave* a cura dei colleghi Tommaso Agnoloni, Lucia Francalanci ed Emilio Sanfilippo, emerge con chiarezza il ruolo di primo piano e il profondo radicamento delle Digital Humanities (DH) e della ricerca in questo ambito condotta al CNR, e in particolare nel DSU con una attività scientifica che si posiziona sia su scala nazionale sia internazionale (come ben rappresentato nei primi due contributi presentati in questo numero dei *DHC Policy Brief*).
L'area progettuale delle Digital Humanities intercetta gli studi e le ricerche nell'ambito delle Scienze Filosofiche, Linguistiche, Giuridiche e Filologico Letterarie che assumono e integrano

metodologicamente le tecnologie digitali per la raccolta e l'acquisizione dei dati (testuali, lessicali, multimediali), la loro rappresentazione ed elaborazione, la gestione e interrogazione[1]. Da qui anche la costituzione del Centro di Digital Humanities del DSU, per rendere visibili e dinamiche le sinergie e la rete di collaborazioni e scambi tra il personale CNR e le istituzioni scientifiche internazionali.

Nel corso degli ultimi anni i network di collaborazione non solo si sono consolidati, ma hanno registrato una crescita esponenziale. Questo dinamico sviluppo è stato fortemente alimentato dal successo nell'ottenimento di progetti finanziati, dall'inserimento strategico e dall'attivazione dei nodi italiani all'interno delle grandi infrastrutture e reti di ricerca europee come CLARIN - *Common Language Resources and Technology Infrastructure*[2], DARIAH - *Digital Research Infrastructure for the Arts and Humanities*[3], E-RIHS - *European Research Infrastructure for Heritage Science*[4], OPERAS - *Open Scholarly Communication in the European Research Area for Social Sciences and Humanities*[5]. Le attività di DH sono condotte infatti anche in sinergia con le infrastrutture di ricerca dell'Ente operanti all'interno delle SSH.

L'attenzione alla dimensione collaborativa della ricerca e alla condivisione dei dati[6] insieme alla partecipazione attiva, propositiva e costante del CNR in questo ambito sono fattori cruciali per questo significativo posizionamento della ricerca in ambito internazionale[7]. Il contributo editoriale, organizzativo e di governance[8] ha permesso al CNR di posizionarsi come un hub centrale del settore delle infrastrutture per le scienze umane e sociali e del patrimonio culturale anche rispetto a piattaforme che nascono dall'impegno congiunto dei nodi italiani delle infrastrutture europee e della collaborazione tra Istituti del DSU come per esempio i progetti FOSSR – *Fostering Open Science in Social Science Research*[9] e H2IOSC – *Humanities and Cultural Heritage open science Cloud*[10], ora confederati in SSHOpenCloud-IT, un ecosistema digitale FAIR per la gestione e condivisione di dati nelle scienze umane e sociali e nel patrimonio culturale.

I risultati di questa prima indagine conoscitiva condotta dal Centro DH del DSU e il patrimonio informativo sinora raccolto rappresentano una base solida, che potrà essere ulteriormente

[1] Cfr. l'insieme delle aree progettuali del DSU: https://www.cnr.it/it/progetti-di-ricerca/area-progettuale/49700/digital-humanities-dus-ad051 (ultima consultazione dei siti web: 5/10/2026).

[2] CLARIN: https://www.clarin.eu/

[3] DARIAH: https://www.dariah.eu/

[4] E-RIHS: https://www.e-rihs.eu/

[5] OPERAS: https://operas-eu.org/

[6] Il CNR ha approvato nel 2023 la *Roadmap per l'Open Science* per la realizzazione e diffusione di pratiche e politiche che garantiscano una ricerca aperta trasparente e condivisa e istituito di recente una unità operativa a supporto. https://sibi.cnr.it/open-science-2/open-science-al-cnr/; https://www.cnr.it/it/news/11966/il-cnr-per-la-scienza-aperta-approvata-la-roadmap-per-la-sua-realizzazione-all-interno-dell-ente

[7] https://www.cnr.it/en/node/23037

[8] https://www.cnr.it/en/european-international-projects

[9] FOSSR: https://www.fossr.eu/

[10] H2IOSC: https://www.h2iosc.cnr.it/

integrata e arricchita per monitorare e valorizzare questa cruciale dimensione della ricerca. Il monitoraggio futuro, infatti, beneficerà dell'inclusione sistematica dei dati provenienti dalle collaborazioni formali e informali attivate anche nei singoli progetti di ricerca, consentendo così di restituire una cartografia ancora più granulare e completa dell'impatto delle Digital Humanities nell'ecosistema scientifico. Il centro CNR DSU - Digital Humanities si propone di fare da raccordo per questo tipo di mappatura al fine anche di poter rappresentare il posizionamento della ricerca nelle DH al CNR nelle diverse costellazioni internazionali.

Figura 1. Mappa dei Centri DH in Europa (elaborata con Gemini Notebook).

La mappa è certamente incompleta perché i diversi data-base sono in continuo e periodico aggiornamento (gran parte dei dati risale a luglio 2025), per realizzarla abbiamo integrato 11 fonti (*NTNU ARTEC, EADH, DARIAH, CLARIN, Aalto University, King's College London, BBAW, EPFL Impresso, GhentCDH, Lund University, DHQ*) ed eliminato sovrapposizioni e doppioni istituzionali, abbiamo così censito e posizionato 43 centri e laboratori unici in 19 paesi europei.

La sfida è però anche quella di aprire il Centro DH del CNR a relazioni dirette con paesi extraeuropei che rispondono ad esigenze linguistiche, storiche e culturali molto specifiche

come, solo per fare un esempio, la *Red de Humanidades Digitales* (RedHD) o il *Centre for Digital Humanities*, University of Lagos (CEDHUL)[11].

2. Reti

La costruzione di reti è un settore strategico per il posizionamento dell'Ente nell'ecosistema della ricerca nazionale e internazionale.

Di cruciale rilevanza per l'integrazione e la diplomazia scientifica dell'Europa mediterranea nelle discipline umanistiche è anche la cooperazione trilaterale tra CNR, CNRS – *Centre National de la Recherche Scientifique*, CSIC – *Consejo Superior de Investigaciones Científicas,* formalizzata e consolidata attraverso due incontri di lavoro istituzionali. Il primo, che si è tenuto a Parigi nel novembre 2024, presso l'*Institut d'études avancées*, intitolato *European research challenges: CNR, CNRS, CSIC Humanities & Social sciences Scientific perspectives,* in cui i tre enti hanno gettato le basi per affrontare congiuntamente le sfide globali dell'era digitale, delle disuguaglianze e della transizione ecologica. Il percorso di convergenza è poi proseguito nel gennaio 2026 a Roma, con il secondo incontro *Exploring shared scientific priorities and fostering inter-institutional collaboration in social science and humanities*, che ha permesso di tradurre le prospettive teoriche in collaborazioni programmatiche concrete, focalizzandosi su aree di cooperazione quali le Digital Humanities, l'impatto delle tecnologie emergenti sull'intelligenza artificiale, l'approccio sociodemografico all'invecchiamento della popolazione e la tutela attiva del patrimonio culturale[12]. Una terza conferenza è prevista per il mese di gennaio 2027 a Madrid, e sarà l'occasione per proseguire il confronto avviato nei precedenti appuntamenti e, soprattutto, per capire come tradurlo.

I tre Enti nell'ambito di questo lavoro di cooperazione scientifica hanno strutturato e attivato tavoli tematici di approfondimento trilaterali. Questa sinergia istituzionale ha mirato a mappare le priorità scientifiche condivise, a favorire la progettazione congiunta in ambito Horizon Europe e a coordinare le risposte alle grandi transizioni globali. Tra questi tavoli uno specifico gruppo di lavoro interistituzionale è stato interamente dedicato alle Digital Humanities e all'Intelligenza Artificiale applicata alle Scienze Umane (Fasano et al. 2025)[13].

Il DSU ha promosso anche recentemente diverse occasioni di confronto scientifico e di riflessione metodologica per sostenere e favorire la collaborazione. Per esempio, i workshop *Exploring shared scientific priorities and fostering inter-institutional collaboration in social science and humanities* e *Workshop on Cultural Intelligence: AI for Social Sciences, Humanities, and Heritage Science* si sono anche focalizzati sul dialogo tra Digital Humanities

[11] REDHD - *Red de Humanidades Digitales* https://humanidadesdigitales.net/; *African Digital Humanities* https://africandh.ku.edu/resources

[12] https://www.dsu.cnr.it/event/european-research-challenges-cnr-cnrs-csic-humanities-social-sciences-scientific-perspectives/; https://www.dsu.cnr.it/event/exploring-shared-scientific-priorities-and-fostering-inter-institutional-collaboration-in-social-science-and-humanities-cnr-cnrs-csic-meeting/

[13] Lo scambio trilaterale ha gemmato anche il progetto "Past discourses, new technological inequalities" finanziato dal programma i-LINK 2025 CSIC Program. Vicepresidencia de Relaciones Internacionales del CSIC, coordinato da Patricia Martín Rodilla dell'IEGPS – *Instituto de Estudios Gallegos Padre Sarmiento*.

e Artificial Intelligence, uno degli asset strategici più complessi dell'attuale riflessione scientifica.

Il framework strategico

Abbiamo pertanto provato a individuare dei pilastri programmatici mirati a consolidare la presenza del CNR nei network globali della conoscenza che massimizzano l'impatto scientifico delle Digital Humanities. Qui di seguito descriviamo i sette pilastri che abbiamo ritenuto strategici a partire dalla riflessione che stiamo conducendo nell'ambito delle DH e del loro posizionamento nell'ecosistema della ricerca (Figure 2 e 3).

Figura 2. Framework strategico (elaborata con Gemini Notebook).

Figura 3. Valorizzazione del personale tecnico-scientifico (elaborata con Gemini Notebook).

1. Valorizzazione del capitale umano

Attraverso programmi di mobilità scientifica e il coinvolgimento attivo del personale di ricerca nei panel di governance della ricerca europea, per sostenerne la presenza anche

rappresentativa nelle associazioni scientifiche nazionali e internazionali di DH e scienze computazionali.

2. Internazionalizzazione
Si tratta di uno dei vettori fondamentali per il DSU del CNR. In questo scenario, un ruolo centrale è occupato dallo sviluppo di alleanze stabili con i principali enti omologhi europei.

3. Infrastrutture di Ricerca
Il rafforzamento dell'interoperabilità e delle sinergie con le grandi infrastrutture di ricerca europee e nazionali di riferimento.

4. Scienza Aperta
La Scienza Aperta è un asset delle DH; per questo è essenziale esplorare l'ecosistema di progetti e infrastrutture di ricerca nati per gestire la complessità e la varietà tipiche di questo settore scientifico. A tal fine sono importanti le sinergie e gli scambi con l'Unità CNR Open Science, con la Biblioteca Centrale CNR "G. Marconi", con l'Osservatorio sui Saperi Umanistici e con l'Osservatorio sulla Scienza Aperta del CNR IRFiS, e con i gruppi di lavoro dell'Ente che operano in questo ambito a diversi livelli (open access, ineguaglianze nell'accesso alle tecnologie, multilinguismo, etc.).

5. Public Engagement e Citizen Science
Attraverso l'attuazione e la promozione di iniziative di *Public Engagement* e *Citizen Science* si connettono le reti di ricerca alle sfide sociali ed economiche del territorio[14].

6. Buone pratiche metodologiche
Attraverso l'utilizzo di strumenti di indirizzo strategico, come i *DHC Policy Brief* promossi dal DSU, si possono sistematizzare e rendere fruibili, riusabili e condivisibili buone pratiche metodologiche.

7. Intelligenza Artificiale
Nella complementarità e intersezione tra DH e Intelligenza Artificiale si giocano molte delle sfide della ricerca collaborativa e di frontiera del futuro. Il DSU è molto attento a queste sinergie, diversi gruppi di ricerca negli Istituti conducono importanti e innovative ricerche nel campo dell'IA e sono attenti ai risvolti etici che queste tecnologie implicano.

Questi pilastri sono interrelati e costituiscono l'architettura sulla quale si sviluppa il framework strategico per creare impatto sulla società scientifica e civile. La dinamica sottesa allo sviluppo del framework genera a sua volta un processo virtuoso di valorizzazione delle competenze

[14] https://www.cnr.it/it/la-ricerca-diffusa

professionali e di ricerca che, se opportunamente sostenute attraverso mobilità e rappresentanza, diventano determinanti per posizionare la ricerca al centro delle scelte istituzionali e politiche a favore della ricerca.

3. Sinergie

Il Centro di Digital Humanities si propone come un nodo strategico e propulsivo nel panorama internazionale delle Digital Humanities per tutti gli Istituti dell'Ente coinvolti in questo settore di ricerca. Attraverso la stipula di accordi di partenariato e la partecipazione attiva a consorzi e reti globali, la struttura intende favorire un dialogo scientifico costante e transdisciplinare. Questa sinergia si tradurrà non solo nella condivisione di buone pratiche e nello scambio metodologico, ma anche nell'adozione di standard condivisi volti a garantire la piena interoperabilità, l'accessibilità e la sostenibilità nel tempo dei dati e delle risorse digitali generati dalla fitta rete di Istituti del DSU del CNR.

Riferimenti bibliografici

Fasano, Gianluca, Cristina Marras, Emilio M. Sanfilippo, Monica Brinzei, Mathieu Husson, et al. 2025. "Unlocking the Potential of DH+AI: Opportunities, Challenges, and Recommendations." *La Lettre de l'InSHS* 4 (92): 48–50.