

Atti del XV Convegno Annuale **AIUCD**

Digitale *e Public Engagement*

Pratiche e prospettive nelle Digital
Humanities

Cagliari 3-4-5 giugno 2026

a cura di

Cristina **Marras** | Andrea **Pergola** | Giampaolo **Salice**

ISBN 9791298618817



Copyright ©2026 AIUCD

Associazione per l'Informatica Umanistica e la Cultura Digitale



Il presente volume e tutti i contributi sono rilasciati sotto licenza Creative Commons Attribution ShareAlike 4.0 International license (CC-BY-SA 4.0). Ogni altro diritto rimane in capo ai singoli autori.

This volume and all contributions are released under the Creative Commons Attribution Share-Alike 4.0 International license (CC-BY-SA 4.0). All other rights retained by the legal owners.

A cura di: Cristina Marras; Andrea Pergola; Giampaolo Salice (2026). *Digitale e Public Engagement: pratiche e prospettive nelle Digital Humanities*, Atti del XV Convegno Annuale AIUCD, Cagliari 3-5 giugno 2026, Università degli Studi di Cagliari.

Ultimo accesso agli URL in data 13 maggio 2026.

Last URL access May, 13 2026.

Si prega di notificare all'editore ogni omissione o errore si riscontri: segreteria [at] aiucd.org

Please notify the publisher of any omissions or errors found: segreteria [at] aiucd.org

Il programma del Convegno AIUCD 2026 è disponibile online:

<https://www.aiucd2026.unica.it/companion/>

The AIUCD 2026 Conference Program is available online:

<https://www.aiucd2026.unica.it/companion/>

I contributi pubblicati nel presente volume hanno ottenuto il parere favorevole da parte di valutatori esperti della materia, attraverso un processo di revisione anonima mediante *double-blind peer review*, effettuata dai membri del Comitato di Programma sotto la supervisione del Comitato Scientifico di AIUCD 2026.

All the papers published in this volume have received favourable reviews by experts in the field of DH, through an anonymous double-blind peer review, carried out by the members of the Programme Committee under the supervision of the Scientific Committee of AIUCD 2026.

Gli Atti del Convegno AIUCD 2026 sono pubblicati come raccolta dei contributi forniti direttamente dagli autori e dalle autrici. I file sono stati raccolti e assemblati senza interventi redazionali significativi da parte dei curatori.

The proceedings of the AIUCD 2026 Conference are published as a collection of contributions provided directly by the authors. The files have been collected and compiled without significant editorial intervention by the editors.

Il logo di AIUCD 2026 è opera di Raffaele Argiolas

The AIUCD 2026 logo was designed by Raffaele Argiolas

La copertina è stata realizzata da Giampaolo Salice

The cover was created by Giampaolo Salice

GENERAL CHAIRS

Giampaolo Salice | Università degli Studi di Cagliari | DH UNICA

Cristina Marras | CNR-ILIESI | AIUCD

COMITATO ORGANIZZATORE / ORGANIZING COMMITTEE

Il Comitato Organizzatore è composto da Giampaolo Salice (Università degli Studi di Cagliari), Cristina Marras (CNR-ILIESI), Raffaele Argiolas, Alessandro Capra, Giommara Carboni, Andrea Pergola, Eleonora Todde (DH UNICA Università degli Studi di Cagliari), Christian D'Agata (Università di Catania), Francesca Frontini (CNR-ILC), Simone Rebora (Università di Verona), Marco Rospocher (Università di Verona), Laura Stochino (ISSASCO), Matteo Tatti e Alice Nozza (Associazione Itzokor).

Iscrizioni / Conference Registration

Eleonora **Todde** – DH UniCa – Università degli Studi di Cagliari

Atti / Conference Proceedings

Andrea **Pergola** – DH UniCa – Università degli Studi di Cagliari

Comunicazione / Communication

Raffaele **Argiolas** – DH UniCa – Università degli Studi di Cagliari

Promozione / Promotion

Giommara **Carboni** – DH UniCa – Università degli Studi di Cagliari

Tecnologo dell'Evento / Event Technologist

Alessandro **Capra** – DH UniCa

Organizzazione locale

Dipartimento di Lettere, Lingue e Beni Culturali, Università degli Studi di Cagliari, Ufficio Amministrativo:
Serena Serra, Giulia Cadoni, Claudia Serri

COMITATO SCIENTIFICO / SCIENTIFIC COMMITTEE

Il Comitato Scientifico del convegno AIUCD 2026 è composto dai Chairs di Programma e dai Chairs delle cinque aree tematiche.

Program Chairs

Giampaolo Salice | Università degli Studi di Cagliari | DH UNICA

Cristina Marras | CNR-ILIESI | AIUCD

DH e co-costruzione del sapere con le comunità: sfide, metodi, strumenti / *DH and co-construction of knowledge with communities: challenges, methods, tools*

Arianna Ciula | King's Digital Lab, London

Greta Franzini | EURAC, Bolzano

Archivi ed edizioni: descrizioni aumentate, accessibilità e sistemi informativi / *Archives and editions: augmented descriptions, accessibility, and information systems*

Federico Valacchi | Università di Macerata

Paolo Monella | Università Kore Enna | AIUCD

Testualità digitali: prospettive, sviluppi e sperimentazioni / *Digital textualities: perspectives, developments, and experimentations*

Simone Rebora | Università di Verona | AIUCD

Alessandro Adamou | Biblioteca Hertziana

Rappresentazione di Dati e Conoscenza / *Data and Knowledge Representation*

Francesca Tomasi | Alma Mater Studiorum – Università di Bologna

Angelo Mario Del Grosso | CNR-ILC

Memorie, storia e patrimoni culturali digitali / *Tangible and Intangible Digital Heritage*

Michela Tardella | CNR-ILIESI

Enrica Salvatori | Università di Pisa

COMITATO DI PROGRAMMA / PROGRAM COMMITTEE

Stefano Allegrezza (Università di Macerata), Laura Antonietti (Université de Versailles Saint Quentin en Yvelines - Université Paris-Saclay), Alessio Antonini (Open University), Liborio P. Barbarino (Università di Catania), Nicola Barbuti (Università di Bari "Aldo Moro"), Stefano Bazzaco (Università di Verona), Andrea Bellandi (CNR-ILC), Giulia Benotto (CNR-ILC), Monica Berti (Leipzig University), Mario A. Bochicchio (Università di Bari "Aldo Moro"), Andrea Bolioli (Ricercatore indipendente), Marco Bombieri (Università di Verona), Flavia Bruni (Università di Chieti-Pescara), Marina Buzzoni (Università Ca' Foscari di Venezia), Alberto Campagnolo (KU Leuven), Vittore Casarosa (CNR-ILC), Raffaele Cioffi (Università di Napoli Federico II), Vincenzo Colaprice (Università di Torino), Francesca Congiu (Università degli Studi di Cagliari), Giuseppe Consolo (Università degli studi di Napoli, Federico II), Christian D'Agata (Università di Catania), Giulia D'Agostino (TU Darmstadt), Elisa D'Argenio (HUN-REN Hungarian Research Centre for Linguistics), Davide Dainese (Alma Mater Studiorum Università di Bologna), Stefano Dall'Aglio (Università Ca' Foscari Venezia), Marilena Daquino (Alma Mater Studiorum Università di Bologna), Mauro De Bari (Università di Bari "Aldo Moro"), Angelo M. Del Grosso (CNR-ILC), Matteo Di Cristofaro (UniMore), Francesca Di Donato (CNR-ILC), Giorgia Di Marcantonio (Università di Macerata), Giorgio Maria Di Nunzio (Università di Padova), Roberto Evangelista (CNR-ISPF), Stefano Ferilli (Università di Bari "Aldo Moro"), Franz Fischer (Università Ca' Foscari di Venezia), Greta H. Franzini (EURAC), Francesca Frontini (CNR-ILC), Daniele Fusi (Università Ca' Foscari di Venezia), Mariangela Giglio (Alma Mater Studiorum Università di Bologna), Tiago Luis Gil (University of Brasilia), Michela Giordano (Università degli Studi di Cagliari), Luca Giovannini (University of Potsdam), Milena Giuffrida (Università di Catania), Giovanna Granata (Università degli Studi di Cagliari), Edmondo Grassi (Università Telematica San Raffaele Roma), Miryam Grasso (Università di Catania), Piergiovanna Grossi (Università di Verona), Marina M. Guglielmi (Università degli Studi di Cagliari), Alessandro Iannella (Università degli Studi di Milano), Sabrina Iorio (Università Napoli Federico II), Michele Lacriola (Università di Siena), Alessandro Laruffa (ISSTOR), Anna Mambelli (UniMore), Francesco Mambrini (Università Cattolica del Sacro Cuore), Francesco Mamei (Università degli Studi di Cagliari), Lorenzo Mancini (CNR-ILIESI), Anna Maria Marras (Università di Torino), Cristina Marras (CNR-ILIESI), Pietro Mazzarisi (Università di Trieste), Barbara McGillivray (King's College London), Federico V. Meschini (Università della Tuscia), Alessio Miaschi (CNR-ILC), Giulia Miglietta (Università del Salento), Paolo Monella (Università Kore di Enna), Rossana Morriello (Università degli Studi di Firenze), Giulia Murgia (Università degli Studi di Cagliari), Enrico Natale (infoclio.ch), Sebastiana Nocco (CNR-ISEM), Serge Noiret (AIPH), Giuseppe Palazzolo (Università di Catania), Mafalda Papini (CNR-ILC), Enrico Pasini (Università di Torino), Giulia Pedonese (CNR-ILC), Paola Peratello (Università Ca' Foscari Venezia), Andrea Pergola (Università degli Studi di Cagliari), Ginevra Peruginelli (CNR-IGSG) University of Groningen, Federico Pianzola (University of Groningen), Fabio C. Pinna (Università degli Studi di Cagliari), Igor Pizzirusso (AIPH), Tiziana Pontillo (Università degli Studi di Cagliari), Valeria Quochi (CNR-ILC), Lisa Reggiani (CNR-ILIESI), Giulia Renda (Alma Mater Studiorum Università di Bologna), Pietro Restaneo (CNR-ILIESI), Dario Rodighiero (University of Groningen), Roberto Rosselli Del Turco (Università di Torino), Federica Rovelli (Università di Pavia), Giampaolo Salice (Università degli Studi di Cagliari), Enrica Salvatori (Università di Pisa), Emilio M. Sanfilippo (CNR-ISTC), Eva Sassolini (CNR-ILC), Manfredi Scanagatta (UniMore), Andrea Schimmenti (Alma Mater Studiorum Università di Bologna), Flavia Sciolette (CNR-ILC), Alessia Scognamiglio (CNR-ISPF), Luigi Serra (CNR-ISEM), Pietro Sichera (CNR-ILIESI), Daria Spampinato (CNR-ISTC), Giulia Speranza (Università di Napoli "L'Orientale"), Rachele Sprugnoli (Università di Parma), Francesco V. Stella (Università di Siena), Timothy Tambassi (Università Ca' Foscari di Venezia), Mirko Tavosanis (Università di Pisa), Francesca Tomasi (Alma Mater Studiorum Università di Bologna), Simona Turbanti (Università di Milano), Nicoletta Usai (Università degli Studi di Cagliari), Marco Venuti (Università di Catania), Gennaro Vessio (Università di Bari "Aldo Moro"), Gabriele Vezzani (Università di Verona), Simone Zenzaro (CNR-ILC).

Enti organizzatori / Organisers

Il XV convegno annuale dell'Associazione per l'Informatica Umanistica e la Cultura Digitale (AIUCD) è organizzata dal DH UNICA - Centro Interdipartimentale per l'Umanistica Digitale e dal Dipartimento di Lettere, Lingue e Beni Culturali dell'Università di Cagliari in collaborazione con AIUCD.



Verso un dizionario narrativo computazionale degli usi linguistici in aree ad alta vulnerabilità sociale

MICHELA BANDINI, SILVIA PICCINI, ANDREA BELLANDI, EMILIANO GIOVANNETTI
CNR-ISTITUTO DI LINGUISTICA COMPUTAZIONALE "A. ZAMPOLLI", ITALY – NAME.SURNAME@ILC.CNR.IT

ABSTRACT (ITALIANO)

Questo contributo presenta la costruzione di un dizionario narrativo computazionale del linguaggio giovanile in contesti di vulnerabilità sociale, sviluppato attraverso un corpus orale di interviste nell'area napoletana nell'ambito di un progetto di Ricerca-Azione. L'analisi integra approcci qualitativi e quantitativi per indagare le strategie definitorie dei parlanti, alla luce della distinzione bruneriana tra pensiero paradigmatico e narrativo. I risultati confluiscono in un dizionario non-standard pensato come una rete concettuale formalizzata secondo il modello OntoLex-Lemon, in linea con i principi FAIR dei Linguistic Linked Open Data.

Parole chiave: dizionario narrativo; linguaggio giovanile; corpus orale; lessico computazionale; OntoLex-Lemon

ABSTRACT (ENGLISH)

Toward a Computational Narrative Dictionary of Language Use in Areas of High Social Vulnerability. This paper presents the construction of a computational narrative dictionary of youth language in contexts of social vulnerability, developed from an oral corpus of interviews run in the Naples area as part of an Action-Research project. The analysis combines qualitative and quantitative approaches to investigate speakers' definitional strategies, in light of Bruner's distinction between paradigmatic and narrative thinking. The results are combined in a non-standard dictionary structured as a conceptual map formalised according to the OntoLex-Lemon model, in line with the FAIR principles of Linguistic Linked Open Data.

Keywords: narrative dictionary; youth language; oral corpus; computational lexicon; OntoLex-Lemon

1. INTRODUZIONE

Il presente contributo illustra le fasi di ideazione e una prima rappresentazione formale di un dizionario narrativo computazionale del linguaggio giovanile in contesti caratterizzati da elevata vulnerabilità sociale. Il lavoro si colloca nell'ambito del progetto sociale *Futuri (im)possibili: diagnosticare il presente e immaginare il futuro con i giovani di Caivano*, promosso dalla Fondazione Rut in collaborazione con la Fondazione Don Calabria per il Sociale, l'Istituto di Linguistica Computazionale "A. Zampolli" (CNR-ILC), l'Istituto di Ricerca su Innovazione e Servizi per lo Sviluppo (CNR-IRISS) e diverse realtà locali. Il dizionario, concepito come uno strumento di supporto alla comprensione dell'ipercomplessità (Morin, 1990) dei fenomeni antro-po-sociali, è pensato per coloro che operano quotidianamente a fianco dei giovani – insegnanti, educatori, operatori sociali, psicologi – ai quali vuole fornire una base di conoscenza utile per progettare futuri interventi educativi e sociali.

L'approccio adottato è quello della Ricerca-Azione, metodo sviluppato dallo psicologo tedesco Kurt Lewin (1946) e ampiamente utilizzato negli studi sociologici in contesti fragili, dove emergono problemi che ostacolano il benessere collettivo. La ricerca non si limita, pertanto, all'osservazione dei fenomeni sociali, ma si configura come un processo dialogico e trasformativo, fondato sull'interazione continua tra teoria e prassi. I dati raccolti orientano l'azione e, al contempo, le azioni producono nuova conoscenza, in un circolo interpretativo dinamico, che si basa su un confronto tra gli attori del contesto di riferimento guidati da una specifica "idea generale" (Elliott, 1993). Proprio in virtù dei destinatari e delle finalità del progetto, il dizionario non è concepito come un elenco statico di significati denotativi, ma include associazioni connotative, interpretazioni ironiche e reti concettuali implicite. Si parte, infatti, dall'assunto che "definire" costituisca un processo cognitivo complesso, che non si riduce a imbrigliare il reale in una formula astratta dai confini rigidi. Nel tentativo di formulare una definizione spesso la mente mobilita reti associative, richiama esperienze vissute, esempi concreti e narrazioni personali. Detto in altri termini, come già sottolineato dallo psicologo statunitense Jerome Bruner (Bruner, 1986; Bruner, 1990), nel

definire entrano in gioco due modalità differenti e complementari che contribuiscono, entrambe, alla costruzione della realtà: da un lato il *pensiero paradigmatico* che tende alla formalizzazione e all'oggettività, e dall'altro il *pensiero narrativo* che fornisce interpretazioni situate, soggettive e culturalmente mediate, organizzate attorno all'esperienza personale e al sé. È proprio il significato narrativo a costituire il tratto distintivo del dizionario che qui presentiamo.

Prima di entrare nel dettaglio della sua struttura e delle scelte di modellazione adottate per la rappresentazione computazionale delle informazioni lessicali, presenteremo brevemente i contesti di raccolta dei dati linguistici, le modalità di acquisizione del materiale e il processo di costruzione del corpus che ha costituito la base empirica di analisi.

2. LA METODOLOGIA DELLA RICERCA: UN CONTRIBUTO LINGUISTICO

In linea con un approccio etnografico-sociolinguistico (Hymes, 1974), la prima fase della ricerca si è concentrata sulla raccolta di dati linguistici. Attraverso tavoli comunitari e consultazioni con educatori, assistenti sociali e operatori del territorio, sono stati individuati tre ambiti di intervento ritenuti rilevanti per osservare pratiche linguistiche spontanee, differenziati in base al grado di strutturazione dei contesti e alle modalità di coinvolgimento dei partecipanti. Nel *contesto strutturato*, rappresentato dall'ambiente scolastico, è stato adottato un approccio controllato, basato sulla somministrazione di questionari cartacei. Gli studenti sono stati coinvolti in attività di riflessione sul significato e sull'uso delle parole attraverso esercizi di comprensione e completamento, domande a scelta multipla e quesiti su scenari ispirati alla vita quotidiana. I *contesti semi-strutturati* hanno incluso ambienti come i doposcuola e i centri diurni, nei quali sono state privilegiate l'osservazione diretta, le interazioni informali e le attività di gruppo pianificate. Infine, i *contesti non strutturati* hanno compreso incontri in presenza con giovani, principalmente in spazi pubblici, nei quali i dati sono stati raccolti mediante questionari digitali e interviste orali registrate.

I contesti strutturato e semi-strutturato, esplorati nell'area di Afragola, hanno svolto una funzione di pre-test metodologico, finalizzata alla messa a punto degli strumenti di raccolta dati da adottare nella fase di test principale che si è svolta a Caivano, in un contesto non strutturato. In questo scenario, i volontari del Centro "Le Salicelle" di Afragola - Istituto Don Calabria, operanti come unità di prossimità (o *unità di strada*), hanno interagito con i giovani del territorio attraverso due principali modalità di raccolta dei dati: *i*) la somministrazione di un questionario digitale tramite *Google Form*; *ii*) la conduzione di interviste semi-strutturate (Adams, 2015), registrate in formato audio o video.

Il questionario digitale¹, diffuso tramite QR code, era composto da domande a risposta chiusa (a scelta multipla) e aperta, articolato in diverse sezioni, volte a raccogliere informazioni specifiche, quali: *i*) dati demografici generali (ad esempio età, genere, luogo di residenza); *ii*) significati attribuiti a una selezione di termini chiave (ad esempio 'capo', 'stupro', 'abuso' e 'ghetto'); *iii*) esempi di uso dei termini in contesti di vita quotidiana e associazioni semantiche. A questo si aggiungevano domande aperte finalizzate a stimolare riflessioni sui significati personali, sulle rappresentazioni culturali e sui potenziali stereotipi associati ai termini chiave.

Le interviste semi-strutturate sono state condotte dai volontari secondo un protocollo composto da sei domande principali, incentrate su tematiche chiave quali il successo, il genere e la legalità. Pur fornendo una base comune di riferimento — a tutti gli intervistati venivano richiesti età, luogo di origine e significato di termini come 'femmina', 'criminale' e 'infame' — il protocollo di intervista prevedeva una certa flessibilità, lasciando agli intervistatori la libertà di adattare o approfondire le domande in base alle esigenze situazionali e alle risposte dei partecipanti. In particolare, gli intervistatori erano incoraggiati a introdurre sollecitazioni esplorative quali "Cosa intendi con questa parola?", "Hai un esempio concreto?", "Conosci qualcuno che per te rappresenta questa idea?", "Dipende dal contesto?" o "Usi questa parola anche con i tuoi amici?", al fine di elicitare chiarimenti, sfumature contestuali ed esempi d'uso situati. Tutte le interviste² sono state raccolte mediante registrazioni audio o, ove possibile, video, previa firma

¹ Ad oggi, sono state raccolte 38 risposte valide da partecipanti di età compresa tra i 12 e i 23 anni, con una netta prevalenza di rispondenti di genere femminile (oltre il 54%), a fronte di una percentuale inferiore al 40% di partecipanti maschi. Dal punto di vista demografico, oltre il 94% dei rispondenti risiede a Caivano, confermando l'efficacia della strategia di R-A territorialmente radicata adottata in questa fase del progetto.

² Tra il 24 aprile e l'8 giugno 2025 sono state complessivamente raccolte 58 interviste, un corpus orale di dimensioni significative. Dal punto di vista demografico, il campione comprende 32 partecipanti di genere femminile e 26 di genere maschile; la maggioranza degli intervistati (58 individui) risiede a Caivano, confermando ulteriormente il forte radicamento territoriale della raccolta dati. Per quanto riguarda la distribuzione anagrafica, i partecipanti hanno un'età

del consenso informato da parte dei partecipanti; nel caso di minori, sono state applicate rigorose misure di tutela della privacy, garantendo che non fossero registrati tratti facciali identificabili³.

3. ANALISI PRELIMINARE DEI DATI: CREAZIONE DI UN DATASET

L'intero processo si è articolato in cinque fasi sequenziali, finalizzate a garantire una rigorosa preparazione del corpus e un'analisi accurata dei dati linguistici.

Fase 1. Le registrazioni audio e video (61 file audio, per un totale di 58 interviste, con una durata complessiva di 2 ore, 18 minuti e 12 secondi) sono state trascritte digitalmente mediante l'ausilio di *NotebookLM* di Google. Una successiva revisione manuale ha consentito di correggere gli errori, standardizzare la formattazione, identificare gli interlocutori e garantire l'anonimizzazione, a tutela della privacy dei partecipanti (vedi nota 3).

Fase 2. Le trascrizioni sono confluite in un corpus orale organizzato in modo sistematico, con pieno allineamento tra file audio/video, metadati e testi trascritti. A ciascun elemento del corpus sono stati assegnati identificatori univoci: le interviste sono state contrassegnate da ID progressivi (ad es. *int01-int58*), i partecipanti da ID corrispondenti (*P01-P58*) e gli intervistatori da identificativi dedicati (*I01-I03*). I materiali sono stati archiviati secondo una rigorosa convenzione di denominazione, che mette in relazione in modo biunivoco registrazioni, trascrizioni e file di metadati in formato JSON, preservando l'integrità di ciascuna intervista⁴.

Fase 3. Una volta strutturato, il corpus ha consentito lo svolgimento di analisi descrittive relative al profilo demografico dei partecipanti e di analisi quantitative sui dati linguistici, utili ad acquisire una panoramica complessiva del dataset, già illustrata nella sezione precedente (note 1 e 2).

Fase 4. È stata quindi condotta un'analisi combinata quantitativa e qualitativa, con particolare attenzione al comportamento terminologico e alle strategie definitorie adottate dai partecipanti. Le analisi sui singoli termini sono state realizzate mediante matrici a doppia entrata, con l'obiettivo di indagare non solo la distribuzione dei significati, ma anche le modalità cognitive attraverso cui i partecipanti hanno espresso il contenuto semantico delle parole (se, ad esempio, attraverso sinonimi o strategie più complesse come la parafrasi esplicativa che enuclea i tratti distintivi del concetto). Le risposte sono state classificate sulla base del quadro binomiale di Bruner illustrato nell'introduzione: le definizioni di tipo scientifico sono state associate ai significati denotativi e interpretate come manifestazioni di pensiero paradigmatico, mentre tutte le altre risposte — incluse le associazioni connotative positive e negative — sono state analizzate come espressioni di pensiero narrativo. Nel loro insieme, le analisi tengono conto in modo integrato dei dati provenienti dai questionari e dalle interviste orali raccolte a Caivano.

Il termine 'femmina' offre un esempio emblematico di tale rappresentazione. Da un lato, i partecipanti hanno attivato una modalità di ragionamento paradigmatica, producendo significati denotativi coerenti con le definizioni repertorate nei dizionari (ad esempio, 'femmina' intesa come "sesso biologico opposto al maschio, atto alla riproduzione"). Tuttavia, nella maggioranza dei casi, attraverso il pensiero narrativo la parola è stata associata a significati connotativi positivi o negativi. I partecipanti hanno fatto spesso ricorso a esperienze autobiografiche, anche mediante esempi prototipici (ad esempio, 'femmina' come "mia madre che fa tutto per me" o associazioni con pratiche quali il vestirsi in modo curato o l'uso del trucco). Il significato è stato organizzato soggettivamente attorno al sé, facendo emergere l'importanza della dimensione culturale: il termine 'femmina' ha dato vita a rappresentazioni socialmente condivise, stereotipizzazioni, giudizi negativi (inferiorità, violenza o sessualizzazione), nonché a narrazioni legate all'identità ("sentirsi donna dentro"). Le associazioni semantiche sono state ricavate direttamente dai testi prodotti dai partecipanti. Gli esempi testuali originali, con le relative frequenze, sono stati conservati e associati alle tipologie definitorie, consentendo analisi qualitative e quantitative. In una fase successiva si

compresa tra i 12 e i 30 anni, con una chiara concentrazione nella fascia 14-17 anni, in linea con il focus primario del progetto sull'uso linguistico adolescenziale.

³ La raccolta e gestione dei dati linguistici è stata condotta nel rispetto del General Data Protection Regulation (GDPR), con particolare attenzione alle specificità del contesto di ricerca, caratterizzato dalla presenza di partecipanti minorenni e situazioni di vulnerabilità sociale. I dati sono stati sottoposti a procedure di anonimizzazione e conservati in ambienti protetti, accessibili esclusivamente ai membri del progetto.

⁴ Il corpus è attualmente conservato in un repository di progetto privato, organizzato per tipologia di dati (audio, trascrizioni, metadati) e documentato tramite un file README. I metadati adottano uno schema personalizzato basato su JSON, sviluppato in funzione delle esigenze specifiche del progetto. Le fasi future prevedono sia il rilascio pubblico del corpus su piattaforme dedicate, come il CLARIN Virtual Language Observatory (VLO), sia la mappatura dei metadati verso standard comunitari per i corpora orali, come Dublin Core, al fine di garantire interoperabilità e riutilizzo a lungo termine.

procederà all'annotazione sistematica del corpus, che permetterà interrogazioni combinate sui dati (cfr. *infra*).

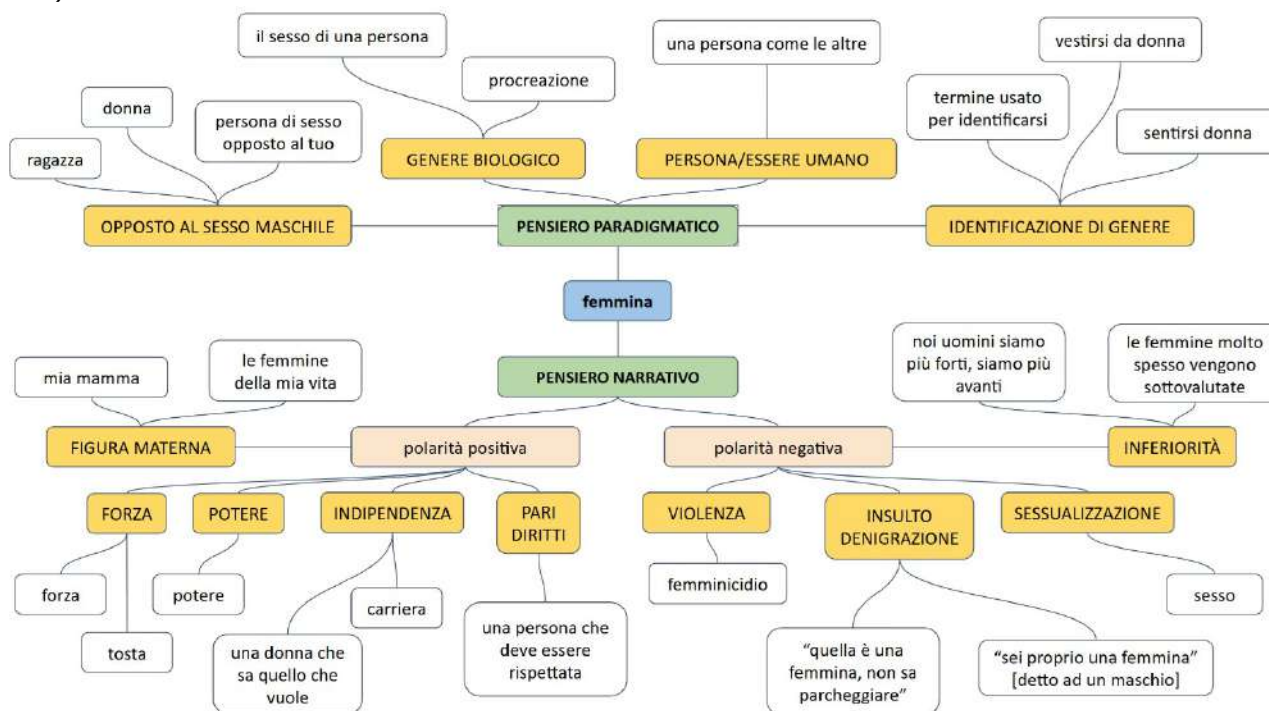


Figura 1. La voce lessicale 'femmina'

Fase 5. Nella fase finale dell'analisi è stato definito il modello di voce lessicografica più adeguato a rappresentare i dati linguistici emersi dalle interviste e dai questionari. Se i dizionari standard tendono a registrare principalmente il significato denotativo delle parole, presentandolo attraverso definizioni paradigmatiche fondate su rapporti di sinonimia e inclusione semantica, un dizionario concepito come strumento di intervento in contesti ad alta vulnerabilità sociale deve invece porre al centro il significato narrativo. Ciò implica dare rilievo alle narrazioni, alle associazioni positive e negative e agli esempi socialmente situati evocati dai giovani intervistati, attraverso i quali il significato si costruisce e si negozia all'interno dell'esperienza vissuta.

Ciascuna voce si articola pertanto in due macro-dimensioni: una paradigmatica e una narrativa, quest'ultima ulteriormente organizzata in base alla polarità (positiva o negativa) e ai concetti evocati dai parlanti. La struttura reticolare che ne deriva, esemplificata in Figura 1, configura le voci come reti di relazioni semantiche.

Al fine di rappresentare e sfruttare appieno questa articolata struttura a "grafo", sia per la consultazione da parte di utenti sia per l'elaborazione automatica dei dati, il dizionario è stato concepito nativamente come risorsa computazionale. La proposta di rappresentazione formale adottata verrà discussa nel paragrafo successivo.

Allo stato attuale, il dizionario comprende 16 voci lessicali, costituite sia da termini monolemmatici sia da espressioni colloquiali o polirematiche ('aborto', 'bastardo', 'capo', 'criminale', 'essere malato', 'essere uno schiaffo', 'fare abuso', 'femmina', 'ghetto', 'guardia', 'infame', 'rispetto', 'si 'na petra', 'sistema', 'soldi', 'stupro'). A ciascuna voce sono associate informazioni morfologiche di base.

4. VERSO UNA RAPPRESENTAZIONE FORMALE DEL LESSICO

La formalizzazione presentata in questa sezione costituisce una prima proposta di modellazione dei dati lessicali emersi dall'analisi del corpus, con particolare attenzione alla rappresentazione del significato narrativo, che rappresenta l'elemento di maggiore originalità del contributo. Al fine di garantire trovabilità, accessibilità, interoperabilità e riuso, in linea con i principi FAIR (Wilkinson et al., 2016), nonché l'integrazione con l'ecosistema dei *Linguistic Linked Open Data* (Cimiano et al., 2020), si è scelto di adottare il modello *OntoLex-Lemon*⁵ (McCrae et al., 2017), ampiamente utilizzato per la

⁵ <https://www.w3.org/2016/05/ontolex/> (cons. 22/04/2026).

rappresentazione di risorse lessicali. La sua flessibilità consente di modellare forme di significato non standard, in particolare quelle legate alla dimensione narrativa, difficilmente rappresentabili attraverso strutture lessicografiche tradizionali. Per la formalizzazione delle parole considerate e dei significati loro attribuiti è stato utilizzato il modulo "core", denominato *ontolex*. La Figura 2 riporta un esempio di formalizzazione, serializzato in formato *Turtle*.

```

01 :fem_entry a ontolex:LexicalEntry ;
02   lexinfo:partOfSpeech lexinfo:noun ;
03   lexinfo:gender lexinfo:feminine ;
04   ontolex:canonicalForm :lex_femmina ;
05   ontolex:sense :fem_sense1, :fem_sense2 .

06 :lex_fe a ontolex:Form ;
07   ontolex:writtenRep "femmina"@it ;
08   lexinfo:number lexinfo:singular .

09 :fem_sense1 a ontolex:LexicalSense ;
10   ontolex:reference :WOMAN ;
11   skos:definition "Essere di sesso opposto a
12     quello maschile, atto a procreare."@it ;
13   dc:type :paradigmatic .

14 :fem_sense2 a ontolex:LexicalSense ;
15   skos:definition "simbolo di forza, potere."@it ;
16   dc:type :narrative .

16 :meaningModalityScheme a skos:ConceptScheme ;
17   skos:prefLabel "Schema delle modalità di costruzione del
18     significato"@it .

18 :paradigmatic a skos:Concept ;
19   skos:prefLabel "paradigmatic"@en , "paradigmatico"@it ;
20   skos:definition "Senso riconducibile a una costruzione
21     paradigmatica del significato: contenuto denotativo e
22     definitorio, allineato a una concettualizzazione
23     lessicografica o scientifica."@it ;
24   skos:inScheme :meaningModalityScheme .

22 :narrative a skos:Concept ;
23   skos:prefLabel "narrative"@en , "narrativo"@it ;
24   skos:definition "Senso riconducibile a una costruzione
25     narrativa del significato: contenuto situato, valutativo
26     ed esperienziale, modellato dal contesto sociale e
27     culturale."@it ;
28   skos:inScheme :meaningModalityScheme .

```

Figura 2. Formalizzazione in OntoLex-Lemon dell'entrata 'femmina': serializzazione in Turtle.

L'entrata 'femmina' è stata rappresentata come istanza della classe *ontolex:LexicalEntry* (riga 01), ossia un'unità astratta che raggruppa le informazioni linguistiche e semantiche associate al termine. L'entrata è descritta tramite proprietà morfosintattiche di base, come la parte del discorso e il genere grammaticale, ed è collegata alla sua forma canonica, modellata come *ontolex:Form* con la corrispondente rappresentazione scritta.

L'aspetto centrale dell'esempio riguarda la rappresentazione dei due tipi di significato introdotti in precedenza, e in particolare la distinzione tra un senso "paradigmatico" e uno "narrativo". Per rappresentarli si è scelto di non estendere il modello *OntoLex-Lemon* (ad esempio introducendo sottoclassi di *ontolex:LexicalSense*), ma di ricorrere a una formalizzazione di questi due elementi come concetti SKOS⁶ (Simple Knowledge Organization System) esterni. Questo approccio riflette la filosofia portante dei *Linked Data*, volta a limitare la proliferazione di entità *ad hoc* e favorire il riuso di ontologie standardizzate.

In dettaglio, i due tipi di senso vengono rappresentati mediante l'utilizzo di due concetti SKOS (*skos:Concept*): "paradigmatic" e "narrative", a loro volta descritti come appartenenti a un *skos:ConceptScheme* che rappresenta l'insieme delle possibili modalità di costruzione del significato (righe 16-17). Attraverso l'utilizzo del vocabolario di metadato Dublin Core Terms⁷, e più specificatamente della proprietà *dct:type*, i tipi vengono assegnati opportunamente ai rispettivi sensi lessicali (righe 12 e 15). La scelta di utilizzare il predicato *dct:type* risiede nella sua versatilità semantica: l'assenza di vincoli di dominio (*domain*) e un intervallo (*range*) esteso a *rdfs:Class* rendono *dct:type* formalmente compatibile con il riferimento a concetti SKOS. Negli sviluppi della ricerca, in ottica di raffinamento semantico, valuteremo la possibilità di introdurre una proprietà dedicata definita come *rdfs:subProperty* di *dct:type*.

Per ragioni di chiarezza espositiva, nell'esempio relativo all'entrata 'femmina' vengono riportati soltanto due sensi, rappresentativi delle due tipologie individuate — uno paradigmatico e uno narrativo — modellati come due istanze di *ontolex:LexicalSense*. Il primo corrisponde al significato standard, denotativo, di 'femmina', come potrebbe essere riportato in un dizionario. Questo senso è stato collegato a un concetto (<WOMAN>⁸), inserito a titolo di esempio, appartenente a un'ontologia di conoscenza condivisa di riferimento. Il secondo senso, invece, formalizza uno degli usi colloquiali individuati

⁶ <https://www.w3.org/TR/skos-reference/> (cons. 22/04/2026).

⁷ <https://www.dublincore.org/specifications/dublin-core/dcmi-terms/> (cons. 22/04/2026).

⁸ I concetti sono indicati tra parentesi uncinate (< >) al fine di distinguerli dai termini lessicali, riportati tra apici singoli (` `).

attraverso le interviste. La descrizione formale di questo tipo di sensi "narrativi", qui presentata in forma preliminare, è attualmente oggetto di ulteriore elaborazione. Sulla base delle analisi condotte, tali sensi non dovranno essere concepiti come semplici accezioni aggiuntive, ma come nodi semantici collegati a concetti lessicali e valutativi specifici, suscettibili di essere qualificati secondo differenti gradi e orientamenti di polarità. Nel caso dell'entrata 'femmina' illustrata nell'esempio, il senso narrativo potrà dunque essere messo in relazione a concetti quali <FORZA> o <POTERE>, attraverso collegamenti caratterizzati da polarità positiva. La formalizzazione di questa componente del modello seguirà le *best practice* dei *Linked Data* e sarà orientata al riuso di vocabolari esistenti, selezionati in funzione della loro adeguatezza a rappresentare dimensioni valutative, concettuali e relazionali del significato.

5. CONCLUSIONI

Il dizionario narrativo degli usi linguistici in aree ad alta vulnerabilità sociale si configura come una risorsa computazionale integrabile nell'ecosistema dei *Linguistic Linked Open Data* e conforme ai principi FAIR, in grado di rendere esplicite le relazioni semantiche tra termini e i molteplici concetti che essi evocano. Le fasi successive della ricerca prevedono un ulteriore affinamento della formalizzazione, attraverso un'annotazione sistematica del corpus e una modellazione più dettagliata delle relazioni semantiche e del loro valore. Parallelamente, i dati verranno resi progressivamente fruibili mediante strumenti di *data visualization* ed esplorazione interattiva. In questa prospettiva, il dizionario si configura come una risorsa operativa e accessibile, a supporto di interventi educativi, sociali e di ricerca interdisciplinare, in linea con i principi della Ricerca-Azione che ne costituiscono il quadro metodologico di riferimento.

BIBLIOGRAFIA

- Adams, W. C. (2015). Conducting semi-structured interviews. In J. S. Wholey, H. P. Hatry & K. E. Newcomer (Ed. 4), *Handbook of Practical Program Evaluation* (pp. 492–505). John Wiley & Sons. <https://doi.org/10.1002/9781119171386.ch19>
- Bruner, J. (1986). *Actual Minds, Possible Worlds*. Harvard University Press. <http://www.jstor.org/stable/j.ctv26070s7>
- Bruner, J. (1990). *Acts of meaning: Four lectures on mind and culture* (Vol. 3). Harvard University Press
- Cimiano, P., Chiarcos, C., McCrae, J. P., & Gracia, J. (2020). *Linguistic linked data*. Springer International Publishing
- Elliott, J. (1993). *La ricerca azione: un quadro di riferimento per l'autovalutazione nelle scuole*. Pozzo G. & Zappi L. (a cura di) Elliott, Giordan, Scurati. *La ricerca azione. Metodiche, strumenti, casi*. Torino, Bollati Boringhieri, 95-120.
- Hymes, Dell. (1974). *Foundations in Sociolinguistics: An Ethnographic Approach*. Philadelphia: University of Pennsylvania Press.
- Lewin, K. (1946). Action research and minority problems. *Journal of social issues*, 2(4), 34-46.
- McCrae, J. P., Bosque-Gil, J., Gracia, J., Buitelaar, P., & Cimiano, P. (2017). The OntoLex-Lemon model: Development and applications. In *Proceedings of eLex 2017*.
- Morin, E. (1990). *Introduction à la pensée complexe*. Paris : Ed. Du Seuil
- Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J., Appleton, G., Axton, M., Baak, A., Blomberg N. et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific data* 3.