

WORKSHOP 2

Verso l'interoperabilità: Linked Open Data e risorse linguistiche per l'italiano

Socie/i proponenti: Claudio Iacobini (Università degli Studi Roma Tre), Eleonora Litta (Università Cattolica del Sacro Cuore), Marco Passarotti (Università Cattolica del Sacro Cuore), Rachele Sprugnoli (Università Cattolica del Sacro Cuore).

Presentazione del workshop

Il workshop ha l'obiettivo di raccogliere contributi relativi all'interoperabilità online tra risorse linguistiche per la lingua italiana, sia di tipo lessicale (ad esempio, lessici, dizionari, thesauri), sia di tipo testuale (ad esempio, corpora). In particolare, sono richiesti contributi che trattino di questioni connesse alla realizzazione e/o allo sfruttamento dell'interoperabilità tra risorse linguistiche dell'italiano tramite l'applicazione dei principii del paradigma Linked Open Data (LOD).

Obiettivi

Macroscopicamente, i contributi attesi si possono ricondurre a due tematiche, non esaustive e non reciprocamente alternative:

(a) Presentazione e/o utilizzo di risorse linguistiche per l'italiano pubblicate, pubblicabili, o ripubblicabili come Linked Open Data. Temi possibili:

- modellizzazione di (meta)dati linguistici;
- (ri)pubblicazione di risorse linguistiche come LOD;
- ricerche linguistiche supportate dal ricorso a (meta)dati forniti da risorse interoperabili in LOD;
- accesso e interrogazione di risorse linguistiche in LOD;
- strumenti a supporto della pubblicazione e/o dell'utilizzo di risorse linguistiche in LOD;
- integrazione di Large Language Model con conoscenza linguistica esplicita, rappresentata in grafi di conoscenza che organizzano (meta)dati provenienti da risorse linguistiche pubblicate come LOD.

(b) Esigenze e critiche relative all'interoperabilità tra risorse linguistiche per l'italiano in LOD. Temi possibili:

- limiti e pregi dell'adozione del paradigma LOD come mezzo a supporto dell'interoperabilità tra risorse linguistiche;
- esigenze di interoperabilità tra risorse riscontrate in lavori di ricerca linguistica;
- iniziative di formazione all'utilizzo di risorse linguistiche in LOD;

- sviluppi futuri dei LOD in ambito linguistico.

Comitato scientifico

Valerio Basile (Università di Torino), Cristina Bosco (Università di Torino), Maria Pia Di Buono (Università degli Studi di Napoli L'Orientale), Francesca Frontini (CNR-Istituto di Linguistica Computazionale, Pisa), Emiliano Giovannetti (CNR-Istituto di Linguistica Computazionale, Pisa), Federica Iurescia (Università Cattolica del Sacro Cuore), Fahad Khan (CNR-Istituto di Linguistica Computazionale, Pisa), Francesco Mambrini (Università Cattolica del Sacro Cuore), Paola Marongiu (CNR-Istituto di Linguistica Computazionale, Pisa), Francesca Murano (Università di Firenze), Matteo Pellegrini (Università Cattolica del Sacro Cuore), Armando Stellato (Università di Roma Tor Vergata), Gaia Tomazzoli (Università La Sapienza), Daniela Vellutino (Università di Salerno).

Lingua dei lavori

Italiano.

Invio delle proposte, tempi e modalità di selezione

Le proposte di contributo, scritte in italiano o in inglese e con una lunghezza massima di 1.000 parole (escluse bibliografia, tabelle e didascalie), dovranno pervenire entro il 20 febbraio 2026 a entrambi i seguenti indirizzi: eleonoramaria.litta@unicatt.it, rachele.sprugnoli@unicatt.it.

L'e-mail, con oggetto "Proposta workshop SLI 2026 LOD", dovrà contenere nome, cognome e affiliazione delle/degli autrici/autori della proposta nonché un indirizzo e-mail a cui si vogliono ricevere tutte le comunicazioni inerenti al workshop. Alla mail andrà allegata la proposta in formato pdf che non dovrà contenere i nomi delle/degli autrici/autori della stessa, perché il processo di revisione sarà in forma anonima.

Il Comitato Scientifico notificherà l'esito della valutazione entro il 31 marzo 2026.

IMPORTANTE: tutte le relatrici e i relatori dovranno essere regolarmente iscritte/i alla Società di Linguistica Italiana al momento dell'inizio del workshop.

Bibliografia

- Bellandi, A. (2022). Le Risorse Linguistiche nell'era del Web Semantico: Un insieme di servizi informatici per la gestione di lessici e terminologie. «AIDAinformazioni», (1-2), 31-50.
- Berners-Lee, T., Hendler, J. & Lassila, O. (2001). The Semantic Web. «Scientific American», 284(5), 28-37.

- Cimiano, P., Chiarcos, C., McCrae, J. P., & Gracia, J. (2020a). Linguistic linked data. Springer International Publishing.
- Cimiano, P., Chiarcos, C., McCrae, J. P., & Gracia, J. (2020b). Linguistic linked open data cloud. In *Linguistic Linked Data: Representation, Generation and Applications* (pp. 29-41). Cham: Springer International Publishing.
- Khan, F., Chiarcos, C., Declerck, T., Di Buono, M. P., Dojchinovski, M., Gracia, J., G.V. Oleskeviciene & Gifu, D. (2022). A survey of guidelines and best practices for the generation, interlinking, publication, and validation of linguistic linked data. In *Proceedings of the 8th Workshop on Linked Data in Linguistics within the 13th Language Resources and Evaluation Conference* (pp. 69-77). Marseille, France. European Language Resources Association.
- Khan, F., & Frontini, F. (2014). Publishing PAROLE SIMPLE CLIPS as linguistic linked open data. In *Proceedings of the First Italian Conference on Computational Linguistics CLiC-it 2014 & and of the Fourth International Workshop EVALITA 2014: 9-11 December 2014, Pisa* (pp. 224-228). Pisa University Press.
- Litta, E., Passarotti, M., Basile, V., Bosco, C., Di Fabio, A., Brasolin, P. (2025). LiITA: a Knowledge Base of Interoperable Resources for Italian. In *Proceedings of the 5th Conference on Language, Data and Knowledge* (pp. 130-135). Naples: UniorPress.
- Passarotti, M., Mambrini, F., Franzini, G., Cecchini, F. M., Litta, E., Moretti, G., Ruffolo, P., & Sprugnoli, R. (2020). Interlinking through lemmas. the lexical collection of the LiLa knowledge base of linguistic resources for Latin. «Studi e Saggi Linguistici», 58(1), 177-212.