

Da LexiCad a LexicHub: note sull'interoperabilità tra risorse lessicografiche

Salvatore Arcidiacono

Università degli Studi di Catania, Italia

Abstract Digital technologies have fostered the emergence of a wide range of specialised lexicographic initiatives, often on a medium or small scale. This paper outlines strategies for enhancing the interoperability of existing platforms, with the aim of integrating information across multiple digital lexical resources. After describing several prototypes developed using the LexiCad system, we propose additional methods for creating an aggregator, named LexicHub, designed to facilitate querying and simultaneous consultation of multiple dictionaries.

Keywords Electronic lexicography. Dictionary writing systems. LexiCad. LexicHub. Interoperability.



Peer review

Submitted 2024-09-23
Accepted 2024-01-07
Published 2025-01-29

Open access

© 2024 Arcidiacono | CC-BY 4.0



Citation Arcidiacono, S. (2024). "Da LexiCad a LexicHub: note sull'interoperabilità tra risorse lessicografiche". *Quaderni Veneti*, 13(1), 165-174.

L'adozione di metodi digitali in lessicografia ha ridotto drasticamente i costi di pubblicazione e abbassato le barriere tecniche ed economiche che disincentivavano l'avvio di nuove iniziative. Ciò ha favorito la nascita di iniziative lessicografiche specializzate, di dimensioni medie e piccole, incentrate su precisi obiettivi di ricerca o intraprese per rispondere alle esigenze di comunità specifiche. Nel caso particolare della lessicografia scientifica e diacronica, questa dinamica si è intrecciata con una tendenza preesistente alla specializzazione, osservabile nel lungo periodo:

All'epoca di Diez e di Meyer-Lübke l'attenzione era rivolta al complesso delle lingue romanze. Presto però, come esemplificato anzitutto dal *Französisches Etymologisches Wörterbuch (FEW)* di Walther von Wartburg, l'interesse cominciò a concentrarsi sulle singole lingue. La fase più recente dello sviluppo si contraddistingue per la concentrazione su singole aree geolinguistiche, singole epoche o singoli settori del lessico (VSES, *DizVenez*, *DI*, ecc.). Rinnunciando all'approccio universale a favore di scelte selettive, si offre la possibilità di raccogliere i rispettivi dati empirici in modo almeno approssimativamente esaustivo e di approfondire la qualità e la precisione delle analisi. Non è da temere un'atomizzazione dei risultati, visto che le opere più specializzate si possono considerare pietre di mosaico che in tappe successive verranno riunite in un quadro d'insieme. (Schweickard 2019, 245)

Il delicato compito di ricomporre questo mosaico informativo viene di solito delegato alle stesse tecnologie che lo hanno prodotto: anche nella lessicografia elettronica è quindi diventato centrale il concetto di interoperabilità, definibile come la capacità di due o più sistemi di comunicare, eseguire programmi o trasferire dati tra varie unità funzionali, senza che lo sviluppatore, né tantomeno l'utente, debba necessariamente conoscere i meccanismi interni di funzionamento di ognuno.¹

Di seguito verranno descritte le strategie di interoperabilità adottate per il sistema lessicografico digitale LexiCad/Pluto, con particolare riferimento a due assi fondamentali: per ciò che concerne le tecnologie, si descriveranno i dispositivi per la condivisione di informazioni e di funzioni; per quanto riguarda gli aspetti concettuali, si individueranno alcuni elementi lessicografici che possono essere utilizzati per il collegamento micro- e macrostrutturale di lessici diversi.

Il sistema LexiCad può essere definito come un framework lessicografico per lo sviluppo di dizionari elettronici. Dal punto di vista tecnologico, LexiCad è assimilabile ai framework utilizzati nello

¹ ISO/IEC 2382:2015(en) *Information Technology Vocabulary*, s.v. *interoperability* (2121317); cf. <https://www.iso.org/obp/ui/en/#iso:std:63598:en>.

sviluppo software, perché accelera la programmazione di nuovi database lessicali e piattaforme informatiche. Da una prospettiva più ampia, invece, LexiCad fornisce un insieme di modelli formalizzati per la lessicografia diacronica, fondati principalmente sui criteri elaborati presso l'istituto del CNR Opera del Vocabolario Italiano (OVI).²

Il software LexiCad è nato nel corso dei lavori preparatori per il *Vocabolario del Siciliano Medievale* (VSM), ma è stato quasi subito adottato dall'OVI per la realizzazione di Pluto (Piattaforma Lessicografica Unica del Tesoro delle Origini), una specifica implementazione di LexiCad sviluppata per il *Tesoro della Lingua Italiana delle Origini* (TLIO).³ L'architettura modulare della piattaforma ha rapidamente reso possibile il riadattamento di LexiCad in contesti di ricerca eterogenei, ma generalmente caratterizzati da un apprezzabile livello di specializzazione.⁴ Lo sviluppo del sistema, non potendosi limitare alle necessità di un singolo progetto, ha dovuto quindi adottare soluzioni valide per un intero 'ecosistema' di risorse, favorendo sistematicamente il riutilizzo delle strutture di dati e delle funzioni.

La capacità di trasferire informazioni e funzioni tra i sistemi della famiglia LexiCad ha quindi costituito un requisito tecnico già dalle prime versioni del *software*. Nell'autunno del 2018, in occasione delle presentazioni di LexiCad/Pluto per il TLIO e della piattaforma PlutoVD per il *Vocabolario Dantesco*, l'interoperabilità tra questi sistemi è stata dimostrata con due applicazioni sperimentali: nella prima, PlutoVD riusciva a integrare automaticamente alcuni nuovi campi nella microstruttura del VD, come quello dedicato all'etimo, compilati automaticamente e in tempo reale a partire da un'interrogazione delle voci archiviate nella base di dati del TLIO; nella seconda, il testo della *Commedia*, pubblicato su un blog esterno, veniva glossato dinamicamente, al clic dell'utente, ricavando i dati dalle voci del VD (Arcidiacono 2019a; 2020a). Entrambi gli esempi evidenziano una tendenza della lessicografia elettronica già riconosciuta da tempo:

² Per una descrizione approfondita si rimanda ad Arcidiacono 2022.

³ Una consistente parte dell'attività di sviluppo di Pluto è oggi condotta nell'ambito del progetto QM (*Quattrocento Meridionale. Il futuro dell'italiano antico* (PRIN 2020).

⁴ LexiCad/Pluto, oltre a essere utilizzato per il trattamento e la riconversione delle voci del TLIO, gestisce attualmente altri quattro progetti lessicografici specializzati: il già citato VSM; il *Vocabolario Dantesco* (VD), il *Vocabolario Latino* (VDL) e il *Vocabolario storico-etimologico del Veneziano* (VEV); è in corso di sviluppo, inoltre, il sistema per il *Dizionario Etimologico e Storico del Napoletano* (DESN). La stessa piattaforma è stata utilizzata per l'*Atlante Grammaticale della Lingua Italiana delle Origini* (AGLIO) e in altri progetti che, a vario titolo, si collocano nell'orizzonte delle *digital humanities*, come nel caso dei progetti ItalArt (*L'italiano delle arti tra Medioevo e Rinascimento*), RdP (*Rime disperse di Petrarca*), la piattaforma del Gruppo Guiron per il *Ciclo di Guiron le Courtois* (<https://guiron.fefonlus.it/>) e la *Bibliografia dei Commenti Danteschi* del CNR-OVI (<http://bcd.ovi.cnr.it>).

it's already clear that the dictionary is morphing from its current incarnation as an autonomous 'product' to something more like a 'service', often embedded in other resources. (Rundell 2012, 29)

Nel primo caso, è il TLIO a selezionare e fornire, 'come servizio', i dati al VD; nel secondo caso, è il VD a occuparsi di disambiguare il lemma attraverso la lettura del contesto, a individuare l'accezione pertinente e a fornire i dati che verranno mostrati al lettore dal blog.

In quello stadio di sviluppo, le applicazioni appena descritte non potevano che limitarsi a mostrare le potenzialità del sistema, perché le due piattaforme non disponevano di una base di dati sufficientemente ampia per fornire dati completi;⁵ malgrado ciò, LexiCad è stato sistematicamente arricchito con una serie di *Application Programming Interface* (API), cioè di interfacce che consentono l'accesso ai dati e ai servizi di una piattaforma non solo agli utenti finali, ma anche ad agenti software esterni, che possono quindi riutilizzare l'informazione. Una prima ricognizione delle API di base è stata fornita in Arcidiacono (2019b): questo studio, sollecitato dall'osservazione delle strategie commerciali di alcuni editori e dai *desiderata* rilevati in un sondaggio condotto da Elexis,⁶ è stato accompagnato da un'attività di sviluppo che ha portato all'attivazione degli *endpoint* nella piattaforma.⁷

Negli anni successivi, l'applicazione delle API di LexiCad, che fornivano i dati nel versatile formato JSON, si è rivelata più utile del previsto per lo sviluppo di funzioni interne alle singole piattaforme, ma le API sono però rimaste sottoutilizzate dal punto di vista dell'interoperabilità tra diverse risorse.

Qualche anno più tardi, le API di LexiCad sono state ulteriormente perfezionate nel tentativo di trasformare le informazioni del TLIO in un dizionario di macchina (inteso come un insieme strutturato di coppie forma-lemma) codificato nel formato XML (Arcidiacono 2022, 149-56). La realizzazione di questo dispositivo è stata sollecitata dalla progettazione delle interfacce di accesso alle voci nei dizionari elettronici e, più nello specifico, da quell'opposizione irriducibile tra i criteri di lemmatizzazione orientati alla facilità di consultazione (che tipicamente accolgono a lemma la forma moderna, o intervengono normalizzando e talvolta ricostruendo la forma del lemma)

5 Ad oggi il processo di raccolta delle informazioni non può dirsi concluso: alla fine del 2023 risultavano pubblicate sul VD 1.341 voci delle 5.545 previste e il sistema Pluto è ancora impegnato nella revisione e nella riconversione delle oltre cinquantamila voci redatte.

6 Dopo l'impostazione delle API documentate in Arcidiacono 2019b, Elexis ha pubblicato un insieme di protocolli condivisi per permettere l'accesso uniforme alle risorse lessicografiche (cf. Kosem et al. 2019; 2021).

7 Un API *endpoint* è l'indirizzo che consente l'interazione con un servizio web tramite una richiesta API.

e un criterio *corpus-based* o, più genericamente, orientato al rispetto delle attestazioni effettive. La mutua esclusività di questi approcci può essere risolta, in un sistema digitale, con il supporto di un dispositivo di lemmatizzazione, come nel caso di LGeRM,⁸ che è stato utilmente applicato nelle risorse lessicografiche, dall'antico francese al francese del Rinascimento. Si è quindi tentato di raccogliere in LexiCad l'informazione necessaria per facilitare l'accesso alle voci con un approccio *thesaurus-based*: quando una ricerca per lemmi non produce nessun risultato, il sistema potrà presentare all'utente una serie di suggerimenti sulla base delle forme registrate nelle voci (punto 0.1 del TLIO).⁹ Grazie a una nuova API, i dati in corso di importazione sul TLIO hanno iniziato a costruire un nuovo dizionario macchina che potrebbe configurarsi come la risorsa di questo tipo più completa per l'italiano delle origini.¹⁰

Gli strumenti fin qui descritti possono essere usati per l'interrogazione trasversale delle risorse sviluppate con LexiCad. Grazie agli *endpoint* esposti dalle piattaforme è possibile costruire un aggregatore, denominato LexicHub, secondo un modello simile a quello della *Stazione lessicografica del Vocabolario Dinamico dell'Italiano moderno* dell'Accademia della Crusca.¹¹

Per fornire un semplice esempio, grazie alla connessione di LexicHub alle API di LexiCad è possibile interrogare simultaneamente il lessionario di due o più dizionari: LexicHub invierà la chiave di ricerca alle API dei diversi vocabolari e raccoglierà i risultati in modo da fornire i collegamenti alle voci da un'interfaccia integrata. Utilizzando gli *endpoint* già funzionanti del VEV e del VSM, per esempio, è possibile visualizzare i risultati della ricerca in due colonne affiancate, come mostrato nella [fig. 1]. La casella di ricerca potrà essere corredata da selettori analoghi a quelli presenti negli indici dei dizionari basati su LexiCad

⁸ *Lemmes, Graphies lemmatisées et Règles Morphologiques*, cf. <http://stella.atilf.fr/LGeRM>.

⁹ L'output dell'*endpoint* ha utilizzato parte della marcatura in XML-TEI delle voci del VSM proposta in Arcidiacono 2013 confrontata con la codifica del lessico *Morphalou* per il francese moderno (<https://repository.ortolang.fr/api/content/morphalou/latest/LISEZ-MOI.html>).

¹⁰ Il *CorpusTLIO* dispone di un dizionario di macchina gestito da GATTO, ma l'utilità di questo strumento è limitata da due fattori: (1) l'attuale versione di GattoWeb non è interoperabile e non espone API; (2) il dizionario di macchina di GATTO si limita a indicizzare le voci attestate nel *Corpus*, mentre le voci del TLIO registrano ulteriori forme morfologiche e varianti grafiche che il redattore è chiamato a individuare al momento della redazione della voce, sia che si tratti di forme sfuggite alla lemmatizzazione, sia di attestazioni fuori corpus (Beltrami 2024, 8-9). Va però osservato che il *Corpus* viene aggiornato periodicamente e il dizionario di macchina in GATTO potrebbe presentare attestazioni assenti nelle voci meno recenti.

¹¹ Cf. <https://www.stazionelessicografica.it/>.

(lemma inizia per / contiene / finisce per,¹² categorie grammaticali,¹³ includi / escludi rinvii, ecc.).¹⁴ Chiaramente, questo metodo può essere utilizzato per l'interrogazione simultanea di altri punti specifici della microstruttura o per l'interrogazione dell'intera voce. Quando l'interrogazione sarà reindirizzata alle definizioni dei vocabolari, si dovrà tenere conto dei criteri redazionali impiegati nei diversi lessici. Per migliorare le interrogazioni, infatti, alcuni dizionari adottano uno stile definitorio che ricorre all'uso strategico di traducenti (d'obbligo, per esempio, nel VSM) e iperonimi (come avviene nel TLIO).¹⁵

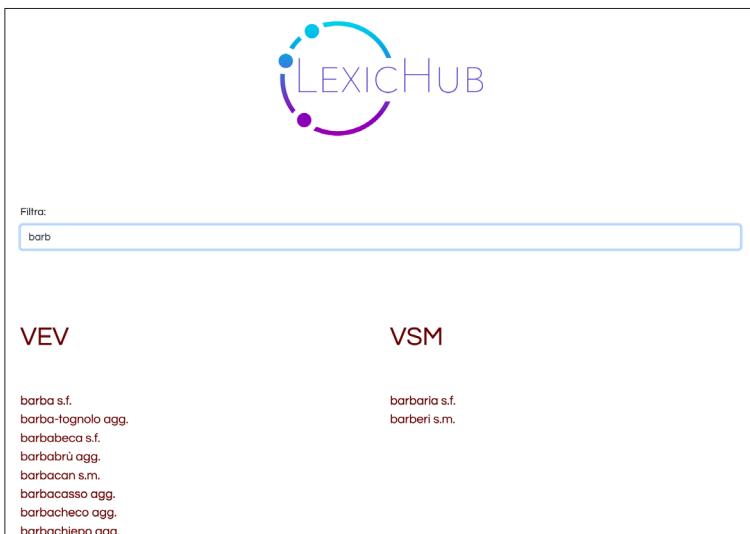


Figura 1 Interrogazione simultanea del VEV e del VSM

I metodi descritti fino a questo momento utilizzano informazioni non strutturate e, per quanto possano essere notevolmente migliorati stabilendo criteri redazionali unificati, difficilmente saranno in grado di costituire strutture adeguatamente formalizzate. Alcuni dizionari basati su LexiCad, come il VSM, il VD e il VDL, dispongono di una sezione in cui vengono registrate le corrispondenze tra i dizionari; queste,

¹² Cf. <http://www.vocabolariodantesco.it/lemmario.php>.

¹³ Cf. <http://www.vocabolariodantescolatino.it/lexicad/lemmario>.

¹⁴ Cf. <http://vev.ovl.cnr.it/lexicad/lemmario>.

¹⁵ Beltrami 2024, 69: «La definizione ideale dovrebbe essere una breve perifrasi sostituibile al lemma nel contesto, e contenente un iperonimo. Con quest'ultimo, si potenziano le ricerche nel vocabolario in rete, mediante la *ricerca nelle definizioni*; per es. la ricerca nelle definizioni di *albero* sarà efficace se tutte le definizioni di nomi di alberi conterranno la parola *albero* nella definizione».

oltre a essere mostrate nell'interfaccia di consultazione come collegamenti ipertestuali, possono essere usate per connettere elementi di risorse diverse attraverso i loro identificativi univoci (come avveniva nel caso già menzionato del collegamento sperimentale tra VD e TLIO).

La relazione tra voce e voce può essere utilmente utilizzata solo per collegare dizionari di una stessa lingua; per il collegamento interlinguistico e interculturale, invece, piuttosto che istituire una connessione tra le voci, in larga parte polisemiche, sarà utile prevedere un collegamento tra le singole definizioni attraverso un elemento di raccordo, come il nodo di una tassonomia o l'entità di un'ontologia. Si tratta, in altre parole, di mappare le definizioni in una struttura onomasiologica condivisa.

Spinto dall'esigenza di organizzare la redazione delle voci su quadri concettuali omogenei, il VSM ha scelto di ricorrere all'impianto onomasiologico dell'*Historical Thesaurus of English* (HTE), un tesastro che organizza il lessico della lingua inglese su base semantica, che ha consentito la classificazione di tutti i significati dell'*Oxford English Dictionary* (OED). In Arcidiacono (2020b), è stata tentata una mappatura sull'HTE di cinquanta lessemi riferibili all'edilizia e alle abitazioni; nello stesso periodo, un tentativo analogo è stato condotto sul TLIO con il progetto DHistOntology, nel quadro di un accordo di collaborazione tra l'OVI e la Real Academia Española.¹⁶ Come osservato da Giuliani e Molina Sangüesa (2020, 344), le mappature sperimentali del VSM e di DHistOntology si accostano all'HTE da prospettive complementari: nel primo caso la tassonomia serve a supportare, con una logica di confronto, la formulazione stessa delle definizioni e l'organizzazione della struttura semantica delle voci; nel secondo caso si tratta di un'analisi 'a posteriori' delle definizioni ricavate dal TLIO e dal DHLE e, di conseguenza, l'accesso alla struttura semantica è sempre mediato dalla definizione lessicografica. In ogni caso, un utilizzo simile di una tassonomia, oltre a dotare il vocabolario di una dimensione onomasiologica, costituisce un potente elemento di raccordo tra banche dati lessicali in rete.

Per concludere, è attualmente in corso la definizione di un modello di voce nel formato *standard* per la codifica del testo XML-TEI, integrato con i moduli forniti da TEI Lex-0.¹⁷ I primi risultati di questa ricerca, pubblicati da Zappalà (2024), riguardano la codifica della microstruttura del TLIO,¹⁸ ma il modello in questione, una volta integra-

¹⁶ I risultati della ricerca sono descritti in Giuliani, Molina Sangüesa 2020.

¹⁷ «TEI Lex-0 is both a technical specification and a set of community-based recommendations for encoding machine-readable dictionaries. It is rooted in the Guidelines of the Text Encoding Initiative (TEI) and delivered as a customization of the TEI schema» (<https://dariah-eric.github.io/lexicalresources/pages/TEILex0/TEILex0.html>).

¹⁸ Questa ricerca è condotta nell'ambito del progetto QM (*Quattrocento Meridionale. Il futuro dell'italiano antico* (PRIN 2020).

to in LexiCad, potrà essere agevolmente trasferito a tutti i dizionari che adottano la piattaforma, favorendo ulteriormente integrazione e l'interoperabilità tra le piattaforme lessicografiche. Lo *standard* tecnico potrà interagire con la struttura concettuale onomasiologica, e, in prospettiva, aprire la strada alla ricerca sulle ontologie,¹⁹ con particolare attenzione al quadro fornito dal modulo Ontolex-Lemon.²⁰

Bibliografia

- AGLIO = Barbato, M.; Faraoni, F. (dir.). *Atlante Grammaticale dell’Italiano delle Origini*.
<http://aglio.ovvi.cnr.it>
- Arcidiacono, S. (2013). «Percorsi di lessicografia computazionale per un Vocabolario del Siciliano Medievale (VSM)». *Bollettino del Centro di studi filologici e linguistici siciliani*, 24, 87-108.
- Arcidiacono, S. (2019a). «Pluto. Piattaforma Lessicografica Unica del Tesoro delle Origini». Leonardi, L.; Squillaciotti, P. (a cura di), *Italiano antico, italiano plurale. Testi e lessico del Medioevo nel mondo digitale = Atti del Convegno Internazionale in occasione delle 40.000 voci del TLIO* (Firenze, 13-14 settembre 2018). Alessandria: Edizioni dell’Orso, 209-17.
- Arcidiacono, S. (2019b). «L’informazione lessicografica nei *Dictionary Writing System* del TLIO». *Bollettino dell’Opera del Vocabolario Italiano*, 24, 381-9.
- Arcidiacono, S. (2020a). «“Forse tu non pensavi ch’io löico fossi!”: metodi computazionali al servizio del VD». Manni, P. (a cura di), *“S’i’ ho ben la parola tua intesa” = Atti della Giornata di presentazione del “Vocabolario Dantesco”* (Firenze, 1 ottobre 2018). Firenze: Accademia della Crusca, 81-92. Quaderni degli Studi di lessicografia italiana 14.
- Arcidiacono, S. (2020b). «Sondaggi sul lessico dell’edilizia in siciliano medievale: il contributo dell’inventario di Alvaro Paternò». *Sinestesieonline*, 9(30), 1-13.
- Arcidiacono, S. (2022). *Lessicografia elettronica e italiano delle origini*. Palermo: Centro di studi filologici e linguistici siciliani.
- Beltrami, P. [1998] (2024). *Norme per la redazione del “Tesoro della Lingua Italiana delle Origini”*. Versione del 28 marzo 2024.
<http://tlio.ovvi.cnr.it/TLIO/NormeTLIO.pdf>
- DESN = De Blasi, N.; Montuori, F. (dir.). *Dizionario Etimologico e Storico del Napoletano*.
- DizVenez = Cortelazzo, M. (2007). *Dizionario veneziano della lingua e della cultura popolare nel XVI secolo*. Padova: La Linea.
- DHLE = Real Academia Española (2013). *Diccionario Histórico de la Lengua Española*. Madrid: Real Academia Española.
<https://www.rae.es/dhle>
- DI = Schweickard, W. (1997-). *Deonomasticon Italicum. Dizionario storico dei derivati da nomi geografici e da nomi di persona*. Tübingen: Niemeyer.

¹⁹ Le ontologie, in ambito informatico, sono strutture formali utilizzate per rappresentare la conoscenza di un dominio specifico, definendo le entità rilevanti e le relazioni che tra queste intercorrono.

²⁰ Cf. <https://www.w3.org/2019/09/lexicog/>.

- FEW = von Wartburg, W. et al. (1922-2002). *Französisches Etymologisches Wörterbuch. Eine Darstellung des galloromanischen Sprachschatzes*, 25 Bde. Bonn; Basel: Klopp; Winter; Teubner; Zbinden.
- Giuliani, M.; Molina Sangüesa, I. (2020). «Hacia una taxonomía integrada en la redacción y revisión de diccionarios históricos». *Bollettino dell'Opera del Vocabolario Italiano*, 25, 325-74.
- Hartmann, R.R.K. (ed.) (2005). «Pure or Hybrid? The Development of Mixed Dictionary Genres». *Facta Universitatis (Linguistics and Literature)*, 3, 193-208.
- HTE = Kay, C. et al. (eds) (2009). *Historical Thesaurus of the Oxford English Dictionary*. 2 vols. Oxford: Oxford University Press.
<https://ht.ac.uk>
- ItalArt = Aresti, A. (dir.) (2020). *L'italiano delle arti tra Medioevo e Rinascimento*.
<http://italart.ovi.cnr.it>
- Kosem, I. et al. (2019). *Early ELEXIS Interoperability Report*.
https://elex.is/wp-content/uploads/2019/08/ELEXIS_D6_1_Early_interoperability_report.pdf
- Kosem, I. et al. (2021). *Intermediate Interoperability Report*.
https://elex.is/wp-content/uploads/2021/02/ELEXIS_D6_3_Intermediate_interoperability_report.pdf
- OED = *Oxford English Dictionary*. Oxford: Oxford University Press. [Supplement 1933; 2nd ed. 1972-86; 3rd ed. 1989].
<http://www.oed.com>
- RdP = Leporatti, R. (dir.). *Le Rime disperse di Petrarca: l'altra faccia del Canzoniere*.
<http://rdp.ovi.cnr.it>
- Rundell, M. (2012). «The Road to Automated Lexicography: An Editor's Viewpoint». Granger, S.; Paquot, M. (eds), *Electronic Lexicography*. Oxford: Oxford University Press.
<https://doi.org/10.1093/acprof:oso/9780199654864.003.0002>
- Schweickard, W. (2019). «Riflessioni sulla lessicografia storica in rete». Leonardi, L.; Squillaciotti, P. (a cura di), *Italiano antico, italiano plurale. Testi e lessico del Medioevo nel mondo digitale = Atti del Convegno Internazionale in occasione delle 40.000 voci del TLIO* (Firenze, 13-14 settembre 2018). Alessandria: Edizioni dell'Orso, 245-6.
- TLIO = *Tesoro della Lingua Italiana delle Origini*. Fondato da P.G. Beltrami, poi diretto da Lino Leonardi, oggi diretto da P. Squillaciotti. Firenze: Opera del Vocabolario Italiano.
<http://tlio.ovi.cnr.it>
- VD = Manni, P.; Leonardi, L. (dir.) (2018-). *Vocabolario Dantesco*.
<http://www.vocabolariodantesco.it>
- VDL = Albanese, G.; Chiesa, P.; Tavoni, M. (dir.) (2021-). *Vocabolario Dantesco Latino*.
<http://www.vocabolariodantescolatino.it>
- VEV = Tomasin, L.; D'Onghia, L. (dir.) (2021-). *Vocabolario storico-etimologico del veneziano*.
<http://vev.ovi.cnr.it>
- VSES = Varvaro, A. (2014). *Vocabolario Storico-Etimologico del Siciliano (VSES)*. Palermo; Strasburgo: Centro di studi filologici e linguistici siciliani; ELiPhi.
- VSM = Pagano, M. (dir.); Arcidiacono, S.; Mosti, R. (vicedir.) (2020-). *Vocabolario del Siciliano Medievale*. Catania: Università di Catania, Centro di studi filologici e linguistici siciliani.
<http://artesia.unict.it/vsm>
- Zappalà, G.L. (2024). «XML-TEI e dizionari born-digital: una proposta per le risorse lessicografiche della rete LexiCad/Pluto». *Me.Te. Digitali. Mediterraneo in rete tra testi e contesti = Atti del XII Convegno AIUCD* (Catania, 28-30 maggio 2024). Catania: AIUCD, 310-17.
<https://doi.org/10.6092/unibo/amsacta/7927>

