

Le risorse della banca dati IS-LeGI per lo studio della lingua del diritto

Francesco Romano[°], Antonio Cammelli[°]

[°]CNR Istituto di informatica giuridica e sistemi giudiziari

For several years, within the framework of institutional research aimed at the documentation of legal language, the Institute of Legal Informatics and Judicial Systems of the CNR has developed an information system (called Semantic Index for the Italian Legal Lexicon, IS-LEGI), used as operational support of the two historical databases: VOCANET-LGI (Italian Legal Language) and LLI (Italian Legislative Language). Significant terms of the selections have been made that are attributed to the prevailing meanings, as well as a conspicuous appropriate phraseology to make better understand the semantic evolution that these entries have suffered during the centuries. Among the terms chosen deserves particular attention to the terms ‘cittadinanza’ and ‘corruzione’ that provides an important contribution for understanding the social changes and cultural influences that have taken place throughout Italy. The system attests the presence of these terms in the legislation, doctrine and practice in a time span ranging from 1377 to 1966.

Keywords: History of legal language - Legal Informatics - Legal linguistics - Semantic tools for information retrieval.

1. La banca dati: descrizione e finalità

L’Istituto di informatica giuridica e sistemi giudiziari (CNR-IGSG) cura l’implementazione di un Indice semantico per il lessico giuridico italiano (IS-LeGI), liberamente consultabile in Rete per mezzo di un intuitivo sistema di interrogazione¹.

L’Indice semantico è stato ideato per consentire un migliore accesso alle banche dati LLI (Lingua Legislativa Italiana) e LGI (Lessico Giuridico Italiano).

¹ Accesso gratuito online

<http://www.ittig.cnr.it/BancheDatiGuide/vgi/islegi/view.status.indice.ragionato.php>

Infatti, la redazione ragionata dell'Indice consente di individuare i collegamenti semantici e concettuali tra i termini giuridici estratti dal corpus degli archivi dell'istituto e di comprendere, non solo la storia, ma anche l'evoluzione degli istituti giuridici che a tale termine fanno riferimento (Fiorelli 2008: 358).

Le unità documentarie che compongono la banca dati sono le schede di cartoncino digitalizzate, ottenute dallo spoglio selettivo dei termini d'interesse giuridico contenuti in circa duemila testi e documenti, tutti pubblicati a stampa, selezionati per rappresentare tanto la varietà delle fonti e dei rami del diritto, quanto la varietà dei secoli e delle regioni in cui la lingua giuridica italiana è stata ed è usata.

Tali testi rappresentano dunque l'uso della lingua del diritto nella legislazione, nella dottrina giuridica e nella prassi, senza dimenticare tipi diversi di documenti, comunque significativi, per l'uso e l'evoluzione storica del lessico giuridico².

Il gruppo di redazione della banca dati (composta da undici persone, tra cui ricercatori e tecnologi del CNR ma anche collaboratori esterni – giuristi, linguisti e tecnologi – il cui supporto risulta imprescindibile per la manutenzione tecnica del sistema e per gli sviluppi dell'interfaccia utente) collabora online, grazie ad un software editoriale. Ciascun redattore, dopo aver effettuato l'accesso alla piattaforma di cooperazione, grazie ad un username e ad una password, è in grado di associare ad un dato lemma, ulteriori rilevanti informazioni che vanno ad arricchire i metadati bibliografici (ad esempio: Schiaffini, Testi fiorentini del Dugento e dei primi del Trecento, Firenze, G. C. Sansoni, tip. L'Impronta, 1954, pp. LVI, 336) già presenti nella banca dati LGI (Romano & Tombesi 2022).

Le informazioni inserite dai redattori sono: le accezioni riferibili al lemma presente nell'unità documentaria, eventuali sintagmi fraseologici relativi al lemma visualizzato e le possibili varianti grafiche visualizzabili nell'unità documentaria (Cammelli & Mariani 2009: 399-408).

Naturalmente, il lavoro di ciascun redattore non diventa immediatamente visibile agli utenti, ma viene pubblicato solo dopo le attività di riallineamento dei due archivi (Back-end e Front-end) che, ovviamente, sono riservate solo a determinati utenti.

La banca dati, è facilmente consultabile grazie ad un sistema di interrogazione che permette di esplorare le parole dell'Indice semantico e i loro significati selezionando direttamente uno dei lemmi catalogati secondo l'ordine alfabetico.

² IS-LeGI attinge dall'archivio costituito dalle oltre 900.000 schede che compongono la banca dati Lessico Giuridico Italiano (LGI).

Dopo aver selezionato un lemma, l'utente può poi leggere le varie accezioni collegate a quel lemma oltre alla «preziosa e variegata fraseologia» di contesto (Lubello 2022: 44). Ovviamente selezionando l'icona "im" è poi possibile visualizzare la scheda immagine in cui il lemma è presente.

Per il futuro è stata già considerata l'ipotesi di «specializzare l'Indice semantico sulla base di un criterio tematico» (Romano, Leggeri & Conti 2022: 197) ed è anche al vaglio la possibilità di introdurre nella maschera di ricerca una apposita funzionalità per la ricerca diretta della fraseologia che, come si vedrà nei paragrafi che seguono, costituisce forse uno dei fattori che caratterizzano questa banca dati.

In passato le risorse della banca dati sono già state usate per documentare l'evoluzione di termini rilevanti per il gergo tecnico-giuridico quali "mandato" (Romano & Cammelli 2018: 519-527) o "codice" (Cammelli & Romano 2020: 401-406), ma anche per documentare la presenza nei documenti del diritto di parole di rilevante impatto sociale: è il caso del termine "razza" (Cammelli *et al.* 2019: 191-200) o della parola "identità" (Cammelli *et al.* 2020: 185-190).

Nel presente contributo saranno invece evidenziati tali informazioni (significati, fraseologia di contesto, varianti grafiche) per altri due importanti termini, quali "corruzione" e "cittadinanza".

2. Le parole 'corruzione' e 'cittadinanza' nei documenti della banca dati

Per fornire utili chiarimenti ed anche al fine di mostrare l'evoluzione (non sempre lineare) della lingua del diritto, abbiamo scelto dal lemmario di IS-LeGI due parole significative: "corruzione" e "cittadinanza".

Come si può facilmente arguire sia l'una che l'altra hanno accompagnato la storia delle nostre comunità, la prima attestandone i lati negativi e, come molte schede documentano, la lotta serrata a questa degenerazione, che ha spesso avvelenato i rapporti delle società del passato come di quelle attuali.

Particolarmente avvertita la corruzione che riguarda l'amministrazione della cosa pubblica.

Sulla cittadinanza la riflessione è molto forte ed attuale, specie nei nostri tempi, che vedono una grande ondata migratoria. Per i migranti la cittadinanza è il simbolo di un nuovo inizio, di speranza di acquisizioni di diritti. La lingua del diritto che ha sempre accompagnato le evoluzioni (ed anche le involuzioni) delle società di riferimento ne è testimone attendibile in quanto fedele.

2.1 Corruzione

Per la parola ‘corruzione’, presente in 59 schede/contesto della banca dati (dal 1377 al 1966) sono state documentate le seguenti accezioni, alle quali vengono associati anche alcuni esempi di sintagmi fraseologici di particolare rilevanza: ‘Compimento o induzione al compimento di atti immorali, libidinosi anche su persona minore di età o socialmente debole; stupro’.

La fraseologia associata a questo significato è quella che segue: agevolare la corruzione, corruzione della figlia minore, corruzione di minorenni, corruzione di minori, corruzione di una vedova, corruzione di una vergine, eccitare la corruzione, favorire la corruzione. Questo primo significato è verificato in documenti che vanno dal 1756 (Bonifazi 1756: 190) al 1966 (Carbone *et al.* 1966: 177).

Un significato ulteriore attestato tra i documenti della banca dati è quello di ‘Danneggiamento, inquinamento’ (corruzione del terreno, corruzione di fonti).

Questo significato è stato individuato in soli 3 documenti, il primo dei quali del 1763 (Guglielmi 1763: 107) e l’ultimo del 1853, attestato all’interno del Regolamento di polizia punitiva pel granducato di Toscana.

In due schede/contesto è stato poi individuato il significato di: ‘Decadenza, perdita dell’esercizio di un diritto’. Si tratta di brani tratti dal Dottor Volgare (De Luca, 1673).

Anche il significato di ‘Degenerazione di una cosa, putrefazione’ è attestato in un documento del 1834 (Liberatore 1834: 128).

Vi è poi un’ultima accezione che è poi quella documentata nel maggior numero di testi (45). L’arco temporale in cui tale significato è attestato va dal 1377, all’interno dello Statuto di Ascoli Piceno, (Zdekauer *et al.* 1910: 117) al 1966 (Carbone *et al.* 1966: 177).

Si tratta del significato che associa il termine a ‘L’indurre, prevalentemente un pubblico ufficiale, con la promessa di denaro o altri vantaggi, a contravvenire ai propri doveri’.

La fraseologia di contesto è quella che segue: avere un titolo di corruzione, colpevole di corruzione, corruzione attiva, corruzione della giustizia, corruzione di funzionario pubblico, corruzione di giudice, corruzione di pubblico ufficiale, corruzione di testimone, delitto di corruzione, denunciare la (per) corruzione, in sequela di corruzione, mezzo di corruzione, non indotto da corruzione, per ragione di corruzione, reato di corruzione, reo di corruzione.

Infine è possibile documentare anche le varianti grafiche di questo termine: corruttione, corrottione, corruzione, corruptione.

2.2 Cittadinanza

Per quanto riguarda il lemma ‘cittadinanza’ le accezioni attestate nelle 199 schede/contesto (dal 1401 al 1965) sono risultate le seguenti:

‘L’insieme della popolazione che costituisce una città, talvolta per estensione una nazione; il complesso dei cittadini’.

La fraseologia collegata è quella che si riporta: incorporazione nella cittadinanza, podestà nella cittadinanza, tranquillità della cittadinanza.

Il significato di ‘Parte della popolazione non nobile ma di rango ragguardevole’ è collegato a sintagmi quali: regia deputazione sulla nobiltà e cittadinanza, stemma di cittadinanza.

Il significato di ‘Stato giuridico derivante dall'appartenenza ad un determinato regno o ad altra autorità costituita, a partire dall'epoca greco-romana’ è invece collegato alla seguente fraseologia: accordare la cittadinanza, acquistare la (una) cittadinanza, acquisto della (di) cittadinanza, aggregare alla cittadinanza, ambire la cittadinanza, ammettere alla cittadinanza, ascritto alla cittadinanza, aspirare alla cittadinanza, assumere la cittadinanza, attestato di cittadinanza, attribuire la cittadinanza, avere la cittadinanza, benefici della cittadinanza, cambiamento di cittadinanza, cambiare cittadinanza, certificato di cittadinanza, cittadinanza accidentale, cittadinanza acquisita iure sanguinis, cittadinanza acquisita iure soli, cittadinanza artificiale, cittadinanza attiva, cittadinanza autenticata, cittadinanza controversa, cittadinanza de intus, cittadinanza de intus et de extra, cittadinanza del marito, cittadinanza della società commerciale, cittadinanza delle persone giuridiche, cittadinanza dello stato, cittadinanza di origine, cittadinanza di primo grado, cittadinanza diocesana, cittadinanza domiciliaria, cittadinanza eminente di primo grado, cittadinanza finta, cittadinanza generale, cittadinanza in forza della legge, cittadinanza legale, cittadinanza mista, cittadinanza naturale, cittadinanza onoraria, cittadinanza originaria, cittadinanza particolare, cittadinanza per atto speciale di grazia, cittadinanza per origine, cittadinanza per privilegio, cittadinanza politica, cittadinanza privilegiata, cittadinanza privilegiativa, cittadinanza romana, cittadinanza straniera, cittadinanza vera e naturale, concedere la cittadinanza, concessione della (di) cittadinanza, condizione di cittadinanza, conferire la cittadinanza, conseguire la cittadinanza, conservare la cittadinanza, controversie concernenti la cittadinanza, dare (la) cittadinanza, decreto di concessione della cittadinanza, determinare la cittadinanza, determinare la cittadinanza del figlio, dichiarazione di abbandono di cittadinanza, dichiarazione di cittadinanza, dichiarazione di eleggere la cittadinanza, dichiarazione relativa alla cittadinanza, diritti di cittadinanza, diritto alla (della, di) cittadinanza, domandare la cittadinanza, eleggere la cittadinanza, essere privilegiato di cittadinanza, godere

del beneficio della cittadinanza, godere il privilegio della cittadinanza, godere la cittadinanza, in ragione pubblica di cittadinanza, istituire una cittadinanza, legge della cittadinanza, legge in materia di cittadinanza, mutamento di cittadinanza, ottenere la cittadinanza, perdere la cittadinanza, perdita della cittadinanza, piccola cittadinanza, prerogativa della (di) cittadinanza, privilegio della (di) cittadinanza, questione di cittadinanza, rapporto di cittadinanza, recuperare la cittadinanza, registro della (di) cittadinanza, relazione personale di cittadinanza, requisito della cittadinanza, riacquistare la cittadinanza, riacquisto della (di) cittadinanza, ricevere la cittadinanza, rinuncia alla cittadinanza, rinunciare alla (la) cittadinanza, seguire la cittadinanza del padre, stato di cittadinanza, strumento di cittadinanza, tassa di cittadinanza, teorica della cittadinanza, togliere cittadinanza, trasmissione della cittadinanza.

Vi è infine un ultimo significato verificato in un documento – I ricordi di Paolo Verzoni del 1653 – della banca dati che è quello di ‘Titolo legale di città e le procedure per il suo ottenimento’ (Petri 1953: 37-84). La fraseologia rilevata è la seguente: bolla della cittadinanza, congregazione sopra la cittadinanza, spezzizione della cittadinanza.

3. Considerazioni conclusive

La banca dati descritta unisce alle possibilità insite negli strumenti tecnologici di conservare e reperire velocemente le informazioni richieste, quella di visionare anche l’immagine del contesto da cui una data parola è tratta, andando dunque a colmare quelle difficoltà che un accesso meramente «mediato dalla tecnologia» può comportare negli studi umanistici (Tomasin 2017: 135).

Ma ci sono anche altri motivi che possono spingere lo storico della lingua o il lessicografo ad usare la banca dati IS-LeGI assieme ad altre banche-dati disponibili online o ai vocabolari storici della lingua italiana.

Il primo fra essi è che «gli archivi storici di riferimento dell’Indice, sebbene riguardino un ambito testuale circoscritto (...) ricoprono un arco temporale molto esteso (X sec. - XX sec.)» permettendo quindi «un ottimo monitoraggio in chiave diacronica delle voci estratte» (Romano & Tombesi 2022: 173).

Si consideri anche che «i maggiori dizionari storici di lingua italiana» hanno una impostazione «fortemente letteraria e quindi meno calibrata sui linguaggi settoriali» (Romano & Tombesi 2022: 173). Tale caratteristica del corpus IS-LeGI ha già permesso in passato e potrebbe farlo anche in futuro, di individuare accezioni nuove o comunque tipiche del lessico giuridico, come verificato, ad esempio, per il lemma “mandato” dove, accanto alle accezioni già «note dai vari

repertori» ne è stata individuata una «nuova tipica del diritto canonico» (Lubello 2022: 43).

Si ricorda anche che la «datazione esibita da ogni scheda-fonte associata a un lemma» può agevolare il lavoro di retrodatazione, così come la possibilità di vedere la «fisionomia originale dei testi a stampa contenuti negli archivi» facilita «analisi linguistiche di vario tipo (studio sulla punteggiatura e sulla sua ratio in diacronia, studio di specifici tratti linguistici)» (Romano & Tombesi 2022: 173).

La banca dati infine può essere usata anche come risorsa da comparare con altre risorse terminologiche (come in passato è già stato sperimentato con la banca dati di dottrina giuridica DoGI) per verificare l'evoluzione del lessico giuridico (Cammelli *et al.* 2022).

Dunque la banca dati IS-LeGI può essere utile per chi studia il lessico giuridico, ma potrebbe risultare particolarmente significativa anche per i giuristi, che hanno a disposizione uno strumento per analizzare le evoluzioni subite nel tempo dai singoli istituti giuridici al fine, ad esempio, di monitorare la trasformazione del sistema nazionale verso il modello comunitario (Marinai *et al.* 2014:181-192).

Inoltre questo corpus annotato è a disposizione di chi voglia comprendere come cambia la nostra società, usando come parametro di riferimento l'evolversi di un linguaggio specialistico, che si affida anche alla lingua comune e che dunque testimonia, nel suo lessico e nella sua sintassi, le trasformazioni delle nostre comunità.

Un esempio molto concreto di questa affermazione è fornito dalla fraseologia associata al significato di cittadinanza, scorrendo la quale è possibile apprezzare una delle contraddizioni irrisolte del nostro tempo e dei nostri ordinamenti e cioè quella tra «l'universalismo dei diritti fondamentali e la cittadinanza» (Ferrajoli 2021: 36). I diritti connessi alla qualifica di cittadino sarebbero (il condizionale è d'obbligo) infatti divenuti universali, come la stessa Carta dell'Onu afferma, ma anche l'armamentario lessicale che si può visualizzare in IS-LeGI ci mostra, invece, quanto gli stessi siano ancora indissolubilmente legati a un mero «accidente di nascita» (Ferrajoli 2021: 36).

Anche per questo motivo tale strumento ultimamente è stato sperimentato con finalità didattiche, non solo all'interno di master dedicati a studenti o operatori del diritto interessati alla storia della lingua del diritto italiano, ma anche nell'ambito di percorsi per le competenze trasversali e l'orientamento per gli studenti delle scuole medie superiori con molteplici scopi, quali approfondire le loro conoscenze giuridiche, linguistiche ma anche storiche, in un momento nel

quale si registra un «problema di diffusa ignoranza e di false idee su eventi del passato specialmente tra i giovani» (Prosperi 2021: 11).

Strumenti interattivi, come quelli qui descritti, oltre a essere perfettamente in sintonia con i più moderni orientamenti in materia di digitalizzazione di beni documentali, associando alla struttura dell'oggetto digitale e al suo contenuto, le informazioni (metadati) che rendono tale struttura e contenuto intellegibili (Barbuti 2022: 22), potrebbero infatti rivelarsi un mezzo efficace per avvicinare i giovani allo studio di una istituzione, vale a dire una delle creazioni oggettivamente utili per un ordinato sviluppo sociale di una determinata comunità di riferimento, che più di ogni altra, va considerata fondamentale «poiché tutte le altre istituzioni presuppongono il linguaggio, mentre il linguaggio non presuppone le altre istituzioni» (Searle 2019: 243).

References

- Barbuti, N. 2022. *La digitalizzazione dei beni documentali*. Milano: Editrice bibliografica.
- Bonifazi, A. 1772. *Nuova succinta pratica civile, e criminale ... distinta in quattro parti, nell'ultima delle quali si aggiunge una istruzione necessarissima per i notari novelli*. Venezia: Giovanni Vitto.
- Cammelli A., Conti S., Peruginelli G. & Romano F. 2022. L'evoluzione storica del linguaggio giuridico: un caso di studio tra due banche dati. In Elena Chiocchetti, Natascia Ralli (a cura di), *Risorse e strumenti per l'elaborazione e la diffusione della terminologia in Italia*, Bolzano: Eurac Research, 186-201.
- Cammelli, A. & Romano, F. 2020. Uso del termine 'codice' nella legislazione, nella dottrina e prassi giuridica. In G. Alfieri, G. Alfonzetti, D. Motta & R. Sardo (eds) *Pragmatica storica dell'italiano. Modelli e usi comunicativi del passato*, Atti del XIII Convegno ASLI - Associazione per la Storia della Lingua Italiana. Firenze: Franco Cesati Editore, 401-406.
- Cammelli, A., Fioravanti, C. & Romano, F. 2020. Identità e linguaggio giuridico. In J. Visconti, M. Manfredini, L. Coveri (eds) *Linguaggi settoriali e specialistici. Sincronia, diacronia, traduzione, variazione*. Atti del XV Congresso SILFI - Società Internazionale di Linguistica e Filologia Italiana. Firenze: Franco Cesati Editore, 185-190.
- Cammelli, A., Fioravanti, C. & Romano, F. 2019. La parola 'razza': analisi diacronica nei testi giuridici antichi e moderni. *Sociologia del diritto*, 2: 191-200.
- Cammelli, A. & Mariani, P. 2009. IS-LeGI. A New On-line Dictionary for a Better Access to the Historical ITTIG Archives Documenting Italian Legal Language. In G. Peruginelli, M. Ragona, (eds). *Law via the Internet. Free Access, Quality of Information, Effectiveness of Rights: Proceedings of the IX International Conference "Law via the Internet"* (Firenze, 30-31 ottobre 2008). Firenze: European Press Academic Publishing, 399-408.
- Carbone, C., Dallari, G., Guglielmi, G., Pescatore, G. & Varvesi, S. 1966. *Formulario del contenzioso amministrativo ... Nuova edizione*. Milano: A. Giuffrè.

- De Luca, G. 1673. *Il dottor volgare, ovvero il compendio di tutta la legge civile, canonica, feudale, e municipale, nelle cose più ricevute in pratica; moralizzato in lingua italiana per istruzione, e comodità maggiore di questa provincia*. Roma: Giuseppe Corvo.
- Ferrajoli, L. 2021. *Perché una Costituzione della Terra?* Torino: Giappichelli.
- Fiorelli, P. 2008. *Intorno alle parole del diritto*. Milano: Guffrè.
- Guglielmi, V. 1763. *Pratica criminale secondo lo stile dello Stato di Toscana*. Pisa: Giovanni Paolo Giovannelli e compagni.
- Liberatore, P. 1834. *Dizionario legale contenente la definizione e la spiegazione dei vocaboli e dei modi di dire usati nell'antica e nuova legislazione canonica civile penale amministrativa secondo l'ultimo suo stato nel Regno delle Due Sicilie*. Napoli: fratelli Rusconi.
- Lubello, S. 2022. *Il diritto da vicino. Intorno ad alcune parole giuridiche dell'italiano*. Alessandria: dell'Orso.
- Marinai, E., Romano, F., Visconti, J. & Dell'Anna, M. V. 2014. *Un corpus di documenti giuridici per il Nuovo Vocabolario dell'Italiano moderno e contemporaneo*. *Informatica e diritto*, XXXIII (1): 181-192.
- Petri, A. (ed) 1952-1953. *Ricordi di Paolo Verzoni*. *Archivio storico pratese*, XXVIII: 37-84.
- Prosperi, A. 2021. *Un tempo senza storia. La distruzione del passato*, Torino: Einaudi.
- Romano, F. & Tombesi, E. 2022. *Documenti giuridici digitali per la storia del diritto e la lessicografia. Italiano digitale*, XXI. <https://id.accademiadellacrusca.org/stampa?id=15817>. (Accesso marzo 2023).
- Romano, F. Leggeri, C. & Conti, S. 2022. *Pluralismo giuridico e storicità del diritto. Per una implementazione dell'Indice Semantico del Lessico Giuridico Italiano*. *Calumet*, 15, 2, pp. 196-207.
- Romano, F. & Cammelli, A. 2018. *Strumenti e metodi per lo studio della lingua del diritto. Le attestazioni del termine «mandato» nei documenti giuridici antichi e contemporanei*. In L. D'Onghia, & L. Tomasin, (eds), *Etimologia e storia delle parole: Atti del XII Convegno ASLI Associazione per la Storia della Lingua Italiana* (Firenze, Accademia della Crusca, 3-5 novembre 2016). Firenze: Franco Cesati Editore, 519-527.
- Tomasin, L. 2017. *L'impronta digitale. Cultura umanistica e tecnologia*. Milano: Carocci Editore.
- Searle, J. R. 2019. *Il mistero della realtà*. Milano: Raffaello Cortina.
- Zdekauer, L. & Sella, P. (ed) 1910. *Statuti di Ascoli Piceno dell'anno MCCCLXXVII*. Roma: Forzani e C.