

## Strumenti e metodi per lo studio in rete della lingua del diritto

Antonio Cammelli, Francesco Romano

*ITTIG CNR, Istituto di Teoria e Tecniche dell'Informazione Giuridica del Consiglio Nazionale delle Ricerche*

**Abstract.** Nell'articolo descriviamo la banca dati denominata Indice semantico per il lessico giuridico italiano (IS-LeGI) che è nata per fornire un accesso migliore a due banche dati da tempo esistenti: VOCANET (un tempo nota come LGI Lingua Giuridica Italiana) e LLI (Lingua Legislativa Italiana). Tale funzionalità è implementata grazie a un software editoriale web based.

**Keywords.** Informatica giuridica, storia linguaggio giuridico

### Introduzione

Tra il 1965 e il 1977 i ricercatori dell'IDG (l'odierno ITTIG CNR) prepararono alcuni archivi lessicali in funzione della redazione di un vocabolario giuridico della lingua italiana (VGI): un archivio integrale elettronico prodotto dalla concordanza cumulativa di cento e più testi e un secondo archivio selettivo, elaborato con mezzi tradizionali e provvisto di un indice unificato per ordine alfabetico di lemmi.

L'impresa scientifica di portare a termine il vocabolario giuridico non è stata completata ma di quella esperienza rimangono importanti lasciti. Il primo è costituito dall'aver implementato un metodo che ha previsto l'uso degli strumenti informatici per il recupero e la conservazione di patrimoni documentari di valore linguistico e storico-giuridico, per la trasmissione più agevole e diffusa della conoscenza e della documentazione dei contenuti e per la gestione di numeri quantitativamente elevati di dati. Inoltre oltre al vasto patrimonio di studi forniti dal gruppo di lavoro IDG-ITTIG che ha lavorato al VGI, sono rimaste consultabili per cittadini e studiosi due banche dati pubbliche, LLI e LGI: un database con oltre duecento testi integrali di testi ufficiali nella loro prima edizione (che coprono un arco temporale che va dal 1539 al 2001) e un archivio digitale con novecentomila schede ottenute dallo spoglio selettivo di testi d'interesse giuridico, pubblica-

ti a stampa (periodo di riferimento dal X al XX secolo).

### Indice semantico online

Attualmente ITTIG [1] prosegue l'implementazione di una banca dati, denominata Indice semantico per il lessico giuridico italiano (IS-LeGI) [2] che è nata per fornire un accesso migliore alle due banche dati sopra rammentate. Tale funzionalità è implementata grazie a un software editoriale web based che permette di associare alcuni metadati ad un dato lemma e alla scheda digitalizzata, che contiene il lemma. Infatti IS-LeGI attinge dall'archivio costituito dalle oltre 900.000 schede che compongono VOCANET (LGI). Alcuni fra questi metadati sono recuperati in via automatica dagli archivi (autore, dati bibliografici dell'opera, area di pertinenza: dottrina [D], prassi [P], legislazione [L], la frequenza assoluta di ricorrenza del lemma negli archivi e quella relativa ai tre settori di riferimento), mentre altri metadati sono creati dal ricercatore per mezzo del software editoriale. Tali metadati aggiuntivi sono: i diversi significati del lemma, che possono variare da contesto a contesto; l'ampia fraseologia giuridica che - meglio di ogni altra informazione - documenta l'evoluzione diacronica di significati [3]: la denominazione Indice semantico nasce proprio dalla natura di questi "metadati aggiuntivi". Infatti questa redazione ragionata consente di individuare i collegamenti semantici e concettuali tra

i termini giuridici estratti dal corpus degli archivi e di comprendere, oltre alla storia di una certa parola, anche l'evoluzione degli istituti giuridici che a tale termine fanno riferimento.

Si prenda come esempio la parola "costituzione", attualmente in fase di redazione. Il lemma è presente in 562 schede della banca dati. Il periodo in cui il lemma è documentato va dal 1250 al 1968 (come detto si può verificare la ripartizione del lemma fra le varie aree di pertinenza: 264 schede per la Dottrina, 184 per la Legislazione e 114 per la Prassi). Il redattore della voce ha, per il momento, individuato 8 significati della parola nella banca dati. Alle accezioni individuate sono poi collegati una serie di sintagmi rilevanti. Alcuni fra essi fanno parte del gergo tecnico giuridico proprio dei tribunali (costituzione del convenuto, costituzione di parte civile, costituzione in cancelleria, costituzione in causa del procuratore, costituzione in termini, costituzione tardiva), altri sono connessi al procedimento penale (costituzione del inquisito, costituzione in carcere, costituzione volontaria), altri al significato di costituzione come norma fondamentale dello Stato (costituzione dello stato, costituzione federativa, costituzione flessibile, costituzione formale, costituzione materiale, costituzione monarchica, costituzione rigida), altri infine sono legati ad istituti giuridici ormai scomparsi (costituzione della dote). Come si vede dal breve excursus esemplificativo la banca dati mette a disposizione una vasta gamma di risorse linguistiche in grado di soddisfare varie esigenze informative.

Le banche dati storico giuridiche dell'ITTIG infatti oltre a essere adoperate in più di un progetto di ricerca, ad esempio per collezionare vaste selezioni di corpora lessicali specialistici [4] potrebbero essere usate all'interno di portali per lo studio della lingua italiana all'estero (si pensi ai recenti progetti ViVit [5] oppure al portale della Lingua Italiana curato dal Ministero degli affari esteri) [6]. Naturalmente tali risorse linguistiche sono già utilizzate per lo studio del lessico giuridico (ad esempio in analisi di testi normativi, [7] non normativi [8] in saggi di storia della comunicazione pubblica [9] o in saggi sul linguaggio burocratico-amministrativo) [10] così come per lo studio della storia del diritto [11]. Più in genera-

le tali risorse potranno essere utili anche in futuro come fonte documentaria per uno studio del presente che dovrà sempre più essere interdisciplinare. Infatti i documenti giuridici costituiscono un punto di vista interessante che potrà fornire "ai futuri giuristi utili elementi per rispondere alle sfide del domani" [12].

La piattaforma IS-LeGI si presta quindi sia alla libera consultazione degli utenti, sia all'apporto collaborativo degli esperti che vogliono partecipare alla redazione on line delle voci sulla base dei documenti forniti dagli archivi.

Tale modalità collaborativa potrà prevedere accessi diversificati per redattori della voce e comitato scientifico, in modo da controllare l'attività svolta (modalità ormai utilizzata in più di un progetto di ricerca [13] in questo settore). Tale intendimento presuppone un allargamento della conoscenza del patrimonio culturale connesso alla scienza giuridica e la Rete in questo senso è ritenuta lo strumento preferenziale per raggiungere questo presupposto [14][15].

### Riferimenti bibliografici

[1] A. Cammelli, P. Mariani, IS-LeGI. A New Online 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" (Florence, 30-31 October 2008), Firenze, European Press Academic Publishing, p. 399, 2009.

[2] [www.ittig.cnr.it/BancheDatiGuide/vgi/islegi/](http://www.ittig.cnr.it/BancheDatiGuide/vgi/islegi/)

[3] A. Cammelli, P. Mariani, op. cit. [1] e anche P. Mariani, IS-LeGI: un dizionario in rete per un migliore accesso al patrimonio giuridico italiano, in *Informatica e diritto*, vol. XVII, fasc. 1-2, pp. 235-244, 2008.

[4] M. V. Dell'Anna, E. Marinai, F. Romano, J. Visconti, Un corpus di testi giuridici per il Nuovo Vocabolario dell'Italiano moderno e contemporaneo: il patrimonio dell'unità ITTIG di Firenze e altre risorse digitali, *Convegno: Piazza delle Lingue*, 2014 - L'italiano elettronico. Vocabo-

lari, corpora, archivi testuali e sonori. Firenze, 6-8 novembre 2014 in Claudio Marazzini, Ludovica Maconi (a cura), "L'italiano elettronico. Vocabolari, corpora, archivi testuali e sonori". ISBN: 978-88-89369-65-4, pp. 223-238 Firenze, Accademia della Crusca, 2016.

[5] [www.viv-it.org/schede/progetto](http://www.viv-it.org/schede/progetto)

[6] [www.linguaitaliana.esteri.it](http://www.linguaitaliana.esteri.it)

[7] L. Spagnolo, *L'italiano costituzionale. Dallo Statuto Albertino alla Costituzione repubblicana*, Loffredo, pp. 412, 2012.

[8] Dell'Anna, M. V., *Aspetti della lingua giuridica. Forme e prassi della scrittura di testi non normativi*, mediAzioni 18, <http://mediazioni.sitlec.unibo.it>, ISSN 1974-4382, p. 8, 2015.

[9] E. Atzori, *La comunicazione pubblica del Comune di Milano. Analisi linguistica (1859-1890)*, Milano, FrancoAngeli, pp. 258, 2009.

[10] S. Covino, F. Marchegiani, *La stampa italiana nell'età napoleonica come collettore e canale di diffusione del linguaggio burocratico-amministrativo e di terminologie speciali: il caso del "Giornale del Trasimeno" (1810-1813)*, in E. Garavelli ed E. Suomela-Härmä (a cura di), *Dal manoscritto al web: canali e modalità di trasmissione dell'italiano. Tecniche, materiali e usi nella storia della lingua*, Atti del XII Congrso SILFI (Società Internazionale di Linguistica e Filologia Italiana), Helsinki, 18-20 giugno 2012, 2 voll., Firenze, Cesati, I, pp. 189-200, 2014.

[11] F. Bambi, *Una nuova lingua per il diritto. Volume I - Il lessico volgare di Andrea Lancia nelle provvisioni fiorentine del 1355-57*, Milano, Giuffrè, pp. 814, 2009.

[12] M. Riberi, *Brevi note sulle fonti storico-giuridiche*, in A. Sciumè, A. A. Cassi (a cura di) *Parole in divenire. Un vademecum per l'uomo occidentale*, Torino, Giappichelli, p. XXV, 2016.

[13] Ad esempio il progetto Papyri.info per la e-

dizione elettronica dei papiri documentari e letterari tramite uno specifico editor ([www.papyri.info/editor/](http://www.papyri.info/editor/)).

[14] D. Weinbergher, *La stanza intelligente: La conoscenza come proprietà della rete*, Torino, Codice Edizioni, pp. 271, 2012.

[15] L. De Biase, *I media civici. Informazione di mutuo soccorso*, Milano Feltrinelli, pp. 200, 2013.

#### Francesco Romano

[francesco.romano@ittig.cnr.it](mailto:francesco.romano@ittig.cnr.it)



Ricercatore dell'Istituto di Teoria e Tecniche dell'Informazione Giuridica del Consiglio nazionale delle ricerche (ITTIG-CNR). Attualmente è responsabile dei progetti CNR: "Strumenti per lo studio del lessico giuridico antico e contemporaneo come patrimonio culturale" e "Controllo di qualità e valutazione dei testi normativi e amministrativi".

#### Antonio Cammelli

[antonio.cammelli@ittig.cnr.it](mailto:antonio.cammelli@ittig.cnr.it)



Dirigente di ricerca del CNR in servizio presso l'Istituto di e Tecniche dell'Informazione Giuridica di Firenze, in quiescenza da alcuni anni, si è sempre occupato dello studio del linguaggio giuridico italiano nella sua declinazione storica, utilizzando il patrimonio lessicale delle banche dati, formatesi in Istituto nel corso di decenni e che costituiscono una preziosa fonte di informazione e studio per gli storici del Diritto. Ideatore del sistema IS-Legi che è un indice intelligente per il recupero delle informazioni delle banche storiche dell'Istituto: L-GI (Lingua Giuridica Italiana) e LLI (Lingua Legislativa Italiana) ora gestibili con un'unica maschera di interrogazione. Insieme al collega Elio Fameli si è occupato della creazione di sistemi di supporto al cittadino soprattutto per il recupero di informazioni in tema di tutela dei diritti. In questo quadro va vista anche la sua collaborazione storica con la commessa romana dell'Istituto. Autore di numerose pubblicazioni al riguardo, consultabili nelle pagine web dell'Istituto, nonché di contributi a seminari e convegni in Italia e all'estero".