

Francesco Romano - Maria Teresa Sagri

Tecnologie per la storia del diritto: gli archivi lessicali storici del Cnr

1. Le banche dati Vocanet e LLI

L'Istituto di teoria e tecniche dell'informazione giuridica del Cnr (Ittig-Cnr) da anni è impegnato in ricerche che riguardano il diritto e le tecnologie della comunicazione e dell'informazione.

L'istituto, che oggi si pone come centro di eccellenza in studi attinenti all'informatica giuridica, deve tuttavia la propria origine all'idea di realizzare un vocabolario storico della lingua giuridica italiana inizialmente nella tradizionale forma a stampa.

Da più di quarant'anni, infatti, gli studiosi del CNR sono impegnati a recuperare e raccogliere un vasto patrimonio lessicale, formando archivi dapprima di schede cartacee, successivamente di testi elaborati elettronicamente con programmi informatici appositamente studiati.

In seguito le nuove opportunità informative offerte dalla rete hanno reso necessario un aggiornamento del progetto originario, soprattutto in merito alla metodologia per la ricerca e consultazione di due archivi lessicali sviluppati per la redazione del vocabolario stesso.

Attualmente sul sito Ittig gli storici del diritto e della lingua hanno a disposizione per le loro ricerche due banche dati di grande interesse.

La prima Vocanet¹ è costituita da circa novecentomila schede cartacee, riprodotte in immagini digitali e collegate a dati lessicali, cronologici e bibliografici, consultabili on line dagli studiosi.

I documenti, che costituiscono uno spoglio selettivo di testi d'interesse giuridico pubblicati a stampa, sono stati scelti in modo da rappresentare sia la varietà delle fonti e la specificità delle discipline giuridiche sia l'evoluzione diacronica del linguaggio giuridico, anche rispetto alle diversità dei contesti geografici.

Le fonti che originano l'archivio si riferiscono a circa duemila testi e documenti di legislazione, dottrina, prassi e altri d'interesse giuridico che vanno dal 960 fino al 1978.

Le schede-fonte sono arricchite da una serie di informazioni quali: il lemma, la data, l'autore e la lingua del testo².

L'interrogazione della banca dati può essere assistita dal programma "VocaNet" e i risultati della ricerca possono essere organizzati in base al lemma (con indicazioni relative alla frequenza totale del termine selezionato per data e autore); all'anno (lemmi presenti in quella data); all'autore (lista alfabetica dei documenti citati e frequenza di lemmi) ed infine alla lingua.

¹ Al sito <http://www.ittig.cnr.it/BancheDatiGuide/vocanet/Index.html> è consultabile gratuitamente la banca dati.

² Le lingue che sono presenti nei diversi documenti della banca dati sono: arabo, austriaco, cinese, danese, ebraico, francese, greco, inglese, indiano, italiano, latino, longobardo, olandese, normanno, norvegese, portoghese, russo, spagnolo, sassone, serbocroato, siamese, somalo, tedesco, turco.

Autore	Lemma	Lingua/Anno	Tipologia	Num.
Breve dei consoli della corte dei mercanti di Pisa	deliberazione	italiano 1311	Legislazione	1 Scheda
Opere di Letteratura, storia, critica	deliberazione	italiano 1335	Legislazione	1 Scheda
Statuto del'Università de' Mercanti e della Corte dell'Università della Città di Siena	deliberazione	italiano 1343	Legislazione	2 Scheda
Opere di Letteratura, storia, critica	deliberazione	italiano 1347	Legislazione	1 Scheda
Statuti della Compagnia dei Dottorali della città di Firenze	deliberazione	italiano 1354	Legislazione	1 Scheda
Giurati (a cura di)	deliberazione	italiano 1356	Prassi	1 Scheda
Ordini della Scapola de' Mercanti fiorentini per la compagnia in Firenze e Anagnone	deliberazione	italiano 1357	Legislazione	1 Scheda
Mari (a cura di)	deliberazione	italiano 1374	Prassi	1 Scheda
Lazzareschi (a cura di)	deliberazione	italiano 1377	Prassi	1 Scheda
Lazzareschi (a cura di)	deliberazione	italiano 1378	Prassi	1 Scheda

Per ciascuna ricerca è possibile visionare il documento digitale con la relativa scheda bibliografica.

L'altro archivio informatizzato ha come oggetto la Lingua Legislativa Italiana (LLI)³ e riguarda Codici, Costituzioni e altre leggi divenute, per vari motivi, fondamentali.

Tali testi sono stati elaborati in modo integrale nella loro prima edizione ufficiale.

I testi selezionati sono 182 e riguardano la legislazione preunitaria che è stata in vigore in Stati di lingua italiana (1539-1921) e una selezione dei testi legislativi più importanti dello Stato italiano.

Archivio LLI (Lingua Legislativa Italiana)

Dati ricercati:
 intero corpus: togoato
 stringa: togoato

Trovate 5 occorrenze in 3 testi, totale records 5 (visualizzati dal numero 1 al numero 5)

Pagine 1

Si LCa (1723)	libro:stolo:z:articolo:z:comma1	del Donante, purché sia Togoato , o 1 Senato, a cui è ...
Si LCa (1723)	libro:stolo:z:articolo:z:comma1	del Decreto del Giudice Togoato , o del Magistrato del Luogo, ...
SC (1834)	parco:stolo:z:sezione:z:paragrafo:z:111	pretiati luogotenenti e di un togoato consigliere: il secondo di un ...
SC (1834)	parco:stolo:z:sezione:z:paragrafo:z:111	uditore del tribunale, di un togoato , uditore della prefettura, ...
Or CI (1920)	pubblica:articolo:z:comma1	e presieduti da un giudice togoato , compongono il Tribunale del ...

ITTIG Motore di ricerca

Attualmente è in corso un'attività di integrazione delle due banche dati, per una più agevole fruizione dei due archivi sopramenzionati.

Da un'unica maschera di ricerca sarà possibile consultare le risorse in un'interrogazione semplice oltre che in modalità ad opzioni di ricerca avanzata.

L'interrogazione simultanea delle due banche dati ne faciliterà la consultazione permettendo al contempo una più agevole ricostruzione dell'evolversi della terminologia giuridica⁴.

³ Al sito <http://www.ittig.cnr.it/BancheDatiGuide/lli/Index.htm> è consultabile gratuitamente la banca dati.

⁴ Questa ristrutturazione delle banche dati è resa possibile anche grazie al contributo dello spin off di Ittig, *Tecnodiritto* (www.tecnodiritto.eu).

A completamento di questa reingegnerizzazione degli archivi informatici esistenti si sta valutando la fattibilità di esportare i dati contenuti in LLI e Vocanet in un formato aperto che possa adottare i descrittori previsti nello standard Dublin Core⁵.

Tale prospettiva è auspicata nell'ottica di condividere i materiali di studio descritti anche all'interno di progetti più vasti⁶.

A tale fine i ricercatori dell'Ittig-Cnr stanno analizzando i metadati che corredano le risorse presenti nelle banche dati LLI e Vocanet al fine poi di verificare se i descrittori dello standard Dublin Core possano essere attribuiti a ciascun metadato.

I testi selezionati sono 182 e riguardano la legislazione preunitaria che è stata in vigore in Stati di lingua italiana (1539-1921) e una selezione dei testi legislativi più importanti dello Stato italiano.

2. La banca dati Is-legi (Indice Semantico per il Lessico Giuridico Italiano)

Una volta reso disponibile in linea il vasto patrimonio documentale raccolto da Ittig, lo sviluppo dell'Indice semantico per il lessico giuridico italiano (IS-LeGI)⁷ ha rappresentato la logica evoluzione delle due banche dati VOCANET e LLI.

L'Indice, attraverso l'apposito programma di interrogazione via web, può essere considerato come un vocabolario giuridico digitale.

La risorsa fornisce infatti alcune chiavi semantiche per migliorare l'utilizzo sinergico dei due Archivi storici dell'Istituto, consentendo l'accesso diretto al corpus e alle voci lessicografiche redatte dai ricercatori Ittig⁸.

Ogni 'voce' contiene la distinzione delle accezioni, i contesti relativi, la fraseologia, le diverse forme e documenta l'evoluzione diacronica di significati⁹.

Ai singoli lemmi sono associati una serie di accezioni, una fraseologia rilevante, eventuali varianti grafiche che documentano la storia del lemma, la frequenza assoluta di ricorrenza del lemma negli archivi e quella relativa ai tre settori di riferimento (dottrina, legislazione, prassi)¹⁰.

L'evoluzione dei significati di un termine permette ai linguisti, così come agli storici del diritto, di comprendere l'evoluzione storica degli istituti giuridici che a tale termine

⁵ Dublin Core (Dublin Core Metadata Element Set - DCMES) è un formato ponte tra più schemi di metadati. È quindi un minimo comune denominatore da utilizzare per integrare risorse e applicazioni eterogenee in servizi di accesso unificato.

⁶ Si pensi a progetti come Europeana, che sta costituendo una biblioteca digitale europea che aggrega contributi già digitalizzati da diverse istituzioni dei paesi dell'Unione Europea, quali libri, film, dipinti, giornali, archivi sonori, mappe, manoscritti ed archivi - <http://www.europeana.eu/portal/>

⁷ Il vocabolario è consultabile gratuitamente tra le banche dati on line dell'ITTIG all'indirizzo <http://www.ittig.cnr.it/BancheDatiGuide/vgi/islegi/>. Inizialmente i due archivi lessicali erano stati creati in vista della redazione di un vocabolario storico dell'italiano giuridico nelle tradizionali forme a stampa.

⁸ <http://www.ittig.cnr.it/BancheDatiGuide/vgi/islegi/Presentazione.htm>

⁹ A. Cammelli, P. Mariani, *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" (Florence, 30-31 October 2008) Firenze, European Press Academic Publishing, 2009 e anche P. Mariani, *IS-LeGI: un dizionario in rete per un migliore accesso al patrimonio giuridico italiano*, in *Informatica e diritto*, 2008, vol. XVII, fasc. 1-2, pp. 235-244.

¹⁰ A. Cammelli, *IS-LeGI in rete: uno strumento di interscambio per lo studio del lessico giuridico italiano*, Firenze, Ittig-Cnr, 2009.

sono connessi.

Questo tipo di indagine risulta particolarmente attuale e significativa anche alla luce delle forti trasformazioni sociali e giuridiche indotte dall'avvento comunitario.

Il fenomeno della globalizzazione culturale rischia infatti di sfumare le differenze esistenti tra gli istituti dei diversi sistemi giuridici e se da un lato diventa auspicabile, nella piena apertura alla modernità, rinsaldare le radici nazionali per non perdere il proprio patrimonio culturale, dall'altro può risultare strategico avere uno strumento per studiare le evoluzioni subite nel tempo dai singoli istituti giuridici al fine di monitorare la trasformazione del sistema nazionale verso il modello comunitario.

Attualmente i ricercatori dell'Ittig-Cnr stanno compilando le voci del vocabolario.

3. Esempi di redazione di voci del vocabolario

Il lavoro editoriale alla base del sistema Is-LeGi, è gestito da un software che permette di associare ai diversi significati del lemma individuato un'ampia fraseologia di contesto e la scheda in cui tale lemma è collocato¹¹, dando così vita ad un dizionario digitale sviluppato in modo collaborativo.

Una volta creata l'accezione il redattore deve collegare l'accezione prescelta al lemma assegnato e possibilmente collegare alla scheda una o più fraseologie ritenute rilevanti nel contesto dato.

Da un apposito form, il software permette infine di inserire le varianti lessicali relative a quel lemma così come attestate nella scheda di riferimento¹².

Volendo analizzare più nello specifico alcuni casi redazionali, il caso del termine "deliberazione" può risultare piuttosto esemplificativo.

Il lemma si presenta nell'archivio con due diverse accezioni documentate da ben 389 schede digitali linkate una ad una con i due diversi sensi.

Nella prima la deliberazione è una "decisione adottata da soggetto che ne ha titolo o da un'autorità, anche collegiale, nonché documento che contiene una decisione adottata" mentre per la seconda accezione, deliberazione è intesa come "l'esame, le considerazioni che precedono una decisione, nonché la volontà di portarla a compimento".

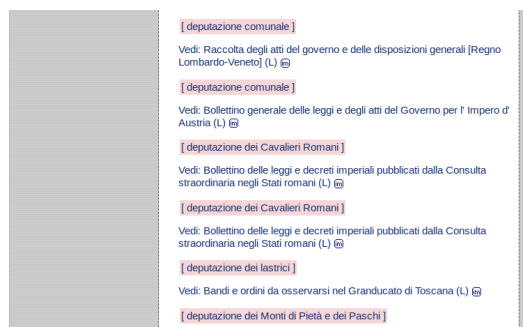
Tali diversi significati sono attestati dai documenti che l'utente può visualizzare *on line*.

Così se per la prima accezione si può vedere come la deliberazione da decisione di un soggetto o di un collegio diventi col tempo anche il documento che contiene quella decisione, la seconda accezione viene documentata come strettamente attinente al diritto penale sia da un'ampia fraseologia (*deliberazione antecedente, deliberazione del delitto, deliberazione dell'omicidio, deliberazione precedente, deliberazione segreta, dolosa deliberazione*) che dagli stessi testi in cui è contenuta (*Ambrosoli: Studi sul Codice penale toscano; Giuliani Giacomo: Principi di giurisprudenza penale;* Nani: *Principi di giurisprudenza criminale; Priori: Pratica criminale; Pasqualigo: Della giurisprudenza criminale*).

Un altro esempio interessante riguarda la voce "deputazione".

¹¹ A. Cammelli, op. cit., p. 5.

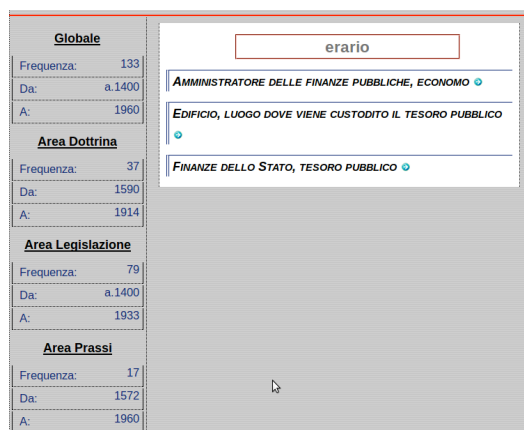
¹² Nel caso di deliberazione sono attestate forme quali: *deliberatione, deliberagione, diliberagione, diliberatione, delliberassione, deliberacione, diliberazione, diliberassione, deliberasione, diliberasione, deliberassione, deliberation*) e per ogni area di riferimento (dottrina, legislazione, prassi) la frequenza del termine e il periodo (nel caso specifico dal 1298 agli anni settanta del Novecento).



In questo caso le accezioni assegnate al lemma (la cui frequenza inizia ad essere attestata dal 1454) testimoniano come al significato di *assegnare un qualcosa* (in ambito giuridico di solito una *causa*) si associ anche l'*affidamento di un incarico o di una missione per un certo scopo*.

Tale incarico istituzionalizzandosi diviene un vero e proprio “*organo o ufficio*” come testimonia l’ampia fraseologia associata a tale terza accezione (*deputazione provinciale, deputazione del Regno, deputazione delle monete, deputazione del (o al) commercio, deputazione di borsa, deputazione del regno, deputazione di sanità*).

Un diverso caso è rappresentato dalla voce “*erario*”.



L’analisi diacronica della documentazione presente nel sistema permette infatti di ricostruire come nel lasso di tempo tra il 1400 e il 1960, sia possibile ricostruire tre diverse accezioni polisemiche.

“*Erario*” secondo il senso di: “*amministratore delle finanze pubbliche*”, ovvero “*economista*”, il secondo senso quale “*edificio, luogo dove viene custodito il tesoro pubblico*” ed infine *erario* secondo il senso di “*finanze dello Stato, tesoro pubblico*”.

Diverso caso è rappresentato dal lemma “*emancipazione*”, *atto giuridico con cui il capofamiglia permette di acquistare ad un figlio o ad altro discendente sottoposto alla sua patria potestà la piena indipendenza giuridica rinunciando nei suoi confronti alla potestà stessa*.

I documenti analizzati permettono di ricostruire come l’istituto tra il 1344 ed il 1965 si trasformi, limitando l’atto di emancipazione alla sola figura del figlio, perdendo così ogni riferimento a qualsiasi altro discendente.

Un altro termine interessante dal quale risulta evidente il modificarsi nel tempo del relativo significato è “*decano*”.

Accanto ad una accezione tutta riconducibile al mondo ecclesiastico in base al quale il

“*decano*” è una vero e proprio titolo¹³ si ha poi una accezione che vede il “*decano*” divenire da “*Membro eminente per anzianità di età o di nomina in una magistratura o in un organismo collegiale*” anche un vero e proprio “*magistrato*” o comunque un “*funzionario pubblico*”.

Anche in questo caso la fraseologia è eloquente attestando forme quali “*decano di anzianità*” o “*decano di età*” ancora in Foramiti, *Enciclopedia legale*, per passare poi al “*decano della (o di) città*” presente in Berti (cur.), *Il processo originale di Galileo Galilei* oppure “*decano della villa*” negli *Statuti de Padova tradotti de latino in vulgare*.

4. Conclusioni

Considerando il valore degli archivi, i cui contenuti sono di estremo interesse per linguisti, storici, documentalisti e storici del diritto, si pensa di creare una rete di esperti lessicografi, che possa partecipare in modo collaborativo alla redazione delle ‘voci’.

In particolare, si vorrebbe aprire la piattaforma IS-LeGI agli esperti che vogliono partecipare alla redazione on line delle voci sulla base dei documenti forniti dai due archivi, prevedendo un sistema di accessi diversificati per la figura del redattore della voce e del valutatore, che convalidi e pubblichi il materiale, sotto la direzione di un comitato scientifico, in uno spazio digitale controllato, ad esempio da un sistema di *peer review*, secondo quella modalità collaborativa ormai utilizzata in più di un progetto di ricerca in questo settore¹⁴.

¹³ Si veda la fraseologia rilevante quale “*cardinale decano*” in Schiappoli, *Manuale del diritto ecclesiastico* o in Lunadoro, *Relatione della Corte di Roma*, o “*decano del (o di) capitolo*”, in *Relazioni dei Rettori veneti in Terraferma* ma anche in Foramiti, *Enciclopedia legale*, oppure “*decano del sacro (s.) collegio*”, in De Luca, *Il dottor volgare* o in Motuproprio, in R. Fantappiè, *Per la storia della diocesi di Prato* e anche in *Notizie di Prato, sì nel temporale, come nello spirituale* e anche in Latini, *Concessione del privilegio di città*.

¹⁴ I. Andorlini “*Edizione e ricostruzione digitale dei testi papiracei*”, relazione al convegno *Diritto Romano e scienze antichistiche nell'era digitale, atti del convegno di studio, Firenze 12 e 13 settembre 2011*. Nella relazione si è illustrato il progetto Papyri.info (<http://www.papyri.info/>) in cui si sta cercando di mettere a punto una edizione elettronica dei papiri documentari e letterari tramite uno specifico editor chiamato Papyrological editor (<http://www.papyri.info/editor/>). Il progetto tende alla edizione on line di testi papiracei e alla integrazione in rete di tutte le risorse disponibili. L'editor, disponibile all'utilizzo anche da parte di una comunità di scienziati esperti in materia e formati all'uso dell'editor per mezzo di seminari formativi, consente anche la funzione “suggerisci modifiche” in una sorta di edizione cooperativa delle risorse documentarie.