

Alcune riflessioni sulle ricerche nelle banche dati brevettuali

Massimo Barbieri

Area Ricerca e Trasferimento Tecnologico - Valorizzazione della Ricerca e Acceleratore d'Impresa - Politecnico di Milano

Come già riportato in altri scritti [1,2], è indubbia l'utilità delle ricerche brevettuali, di cui cercherò di affrontare alcuni aspetti nei paragrafi seguenti.

1. "Capire" l'invenzione

Non sempre è facile, ma per chi svolge ricerche brevettuali è importante comprendere quali sono gli elementi essenziali e caratteristici di un'invenzione per poter individuare un set di parole chiave (e sinonimi) corrispondenti e nel migliore dei casi uno o più simboli di classificazione o codici di indicizzazione.

2. "Non esaustività" delle ricerche

È noto che le ricerche sullo stato dell'arte brevettuale non possono essere complete al 100% [3], a causa di fattori quali l'inevitabile inadeguatezza di qualsiasi sistema di recupero dell'informazione¹.

3. Scelta del database

La scelta del database dipende essenzialmente dal tipo di informazione da ricercare: per esempio, la ricerca di composti, reazioni chimiche o strutture di tipo "Markush" è possibile solo su database commerciali (Dialog, STN, Questel-Orbit) e non su quelli gratuiti tipo Esp@cenet.

¹ Durante il 19° secolo i brevetti potevano essere ricercati solo conoscendo il sistema di classificazione degli uffici brevetti nazionali e andando personalmente nelle sedi ad effettuare ricerche manuali. Solo lo sviluppo della tecnologia informatica ha reso possibile un incremento nel recupero dell'informazione. [E. S. Simmons - "PATENT DATABASES AND GRESHAM'S LAW" - World Patent Information (2006), 28, p. 291 - 293

Una banca dati "ideale" dovrebbe contenere tutte le domande, i brevetti concessi, nonché i modelli d'utilità di tutti gli uffici brevetti nazionali ed internazionali. Dovrebbero, inoltre, essere disponibili i dati bibliografici, il testo completo ed i disegni, nonché lo stato legale per ciascun brevetto[4]. Infine, i brevetti dovrebbero essere tradotti in un'unica lingua, in modo da facilitare la ricerca e la valutazione di ciò che realmente fa parte dello stato della tecnica.

4. Parole chiave o classificazione?

Per le ricerche più semplici può essere sufficiente una combinazione di parole chiavi, mentre per quelle più complesse occorre individuare uno o più codici di classificazione in cui rientra l'invenzione.

Una ricerca in una classe (IPC) ampia raramente è utile, a causa del numero elevato di risultati ottenuti. Inoltre, la classificazione IPC è applicata da diversi uffici brevetti nazionali e regionali ed è pertanto vincolata all'interpretazione degli esaminatori².

In questo caso è preferibile utilizzare i codici ECLA³, che sono non solo più specifici rispetto all'IPC [si

² I codici ECLA sono attribuiti solo dall'Ufficio Europeo Brevetti.

³ Lo schema della classificazione ECLA è basato sull'ultima edizione della classificazione internazionale, a cui sono aggiunte sottodivisioni e testo. Il testo aggiuntivo è aggiunto in parentesi ed è distinguibile perché preceduto dalla lettera N (che significa "Non ufficiale": il testo ufficiale rimane comunque quello della classificazione internazionale) e da codici del tipo [N9809] (che significa "Nuovo": data d'introduzione settembre 1998) o [C9601] (che significa "Modificato" nel gennaio 1996). Le regole per decidere le sottodivisioni sono piuttosto semplici: l'ordine è alfabetico, una lettera è seguita da un numero e poi da un'altra lettera, mentre il punto s'inserisce quando al sottogruppo viene aggiunta una lettera oppure un numero. L'esaminatore decide la selezione della lettera (che a volte è tratta dal testo aggiuntivo) e la sottodivisione.

Una recente novità in Esp@cenet riguarda la modalità in cui i codici ECLA sono ricercabili. Se si inserisce l'asterisco come ultimo carattere alla fine del codice, è possibile recuperare tutti i documenti contenuti in eventuali sottodivisioni [S. van Dulken -"NEWS

tratta di una versione più dettagliata (tabella 1) della classificazione internazionale], ma anche dinamici, ovvero sono costantemente sottoposti a revisione per far fronte ai continui sviluppi della tecnologia.

<i>Classificazione</i>	<i>Numero di gruppi</i>
IPC	70.000
ECLA	130.000
JPO: F-terms	180.000
USPC	160.000

Tabella 1 - Sistemi di classificazione

Uno degli svantaggi nell'uso delle classi ECLA riguarda il fatto che, ad esempio, ad alcune categorie di domande di brevetto, quali quelle depositate negli Stati dell'estremo oriente (Giappone, Corea, Cina, ecc...) - a meno che non siano estese a livello internazionale - e spesso tutte quelle pubblicate nell'ultimo anno, non viene attribuito alcun codice ECLA⁴ [5]. Una ricerca brevettuale eseguita solo sulle classi ECLA non può certamente definirsi completa⁵, anche nel caso in cui sia

ON PATENT, TRADEMARK AND DESIGN DATABASES ON THE INTERNET" - World Patent Information (2006), 28, p. 348 - 349].

⁴ Da un documento dell'UEB risulta che, per esempio, i documenti brevettuali giapponesi presenti in DocDB (master database, che contiene i dati bibliografici, i riassunti e le citazioni brevettuali) sono 13,6 M, mentre quelli classificati in ECLA sono 2,4 M (18%). Ciò significa che i brevetti giapponesi non sono classificati in modo sistematico e che, pertanto, la ricerca in ECLA non è sufficiente ad ottenere un completo recupero dell'informazione brevettuale. [Heiko Wongel - "IPC8 AND ECLA - COMPREHENSIVE SEARCH AND RETRIEVAL" - <http://academy.epo.org/schedule/2006/ic04/IC-04/IPC%20and%20ECLA/player.html>

⁵ Con la ricerca eseguita solo con le classi ECLA si perdono circa 1/3 dei documenti. [Eugenio Archontopoulos - "COMPUTER PROGRAMS: FROM DESIGN TO RUNTIME"

- <http://academy.epo.org/schedule/2006/ic04/IC-04/S18p.pdf>]

individuata la classe corrispondente all'invenzione in esame.

La classificazione Derwent può essere molto utile in certi settori tecnologici (chimico, farmaceutico), perché caratterizzata da un'ampia indicizzazione [6].

La classificazione ICO⁶ (In Computer Only) è riportata solo in poche banche dati (tra cui QPAT), perché non viene utilizzata in tutti i settori della tecnica e soprattutto perché non è obbligatoria.

I codici di classificazione ICO riguardano sia gli aspetti non inventivi sia le caratteristiche aggiuntive (dettagli, applicazioni, materiali, ecc...) di un'invenzione.

Tali codici possono avere lo stesso significato⁷ del corrispondente codice ECLA, possono costituire una sottodivisione, essere del tutto indipendenti⁸ e infine una combinazione di più simboli.

La classificazione statunitense⁹, costituita da classi e sottoclassi, è basata sulla tecnologia descritta nei brevetti US. Le classi sono definite in base a diversi fattori: utilità, struttura e tipo d'industria [7, 8]. Uno degli svantaggi riguarda il fatto che solo i

⁶ Si tratta di una classificazione che si sviluppa secondo uno schema parallelo all'ECLA (con una modifica della lettera che definisce la sezione: alla A si sostituirà, per esempio, la lettera K, ecc...), contenente informazioni più dettagliate e supplementari, ma caratterizzata da una riclassificazione meno sistematica ed effettuata durante la ricerca [Roberto Iasevoli - "INTRODUCTION TO THE EUROPEAN PATENT CLASSIFICATION"

- http://academy.epo.org/schedule/e_learning/ecla/player.html].

⁷ In questo caso sono specificati gli aspetti non inventivi.

⁸ In entrambi i casi non si ha un corrispondente codice ECLA

⁹ Entrata in vigore nel 1836, subì una prima revisione approssimativamente nel 1898. Le revisioni sono effettuate ogni due mesi ed aggiornate nelle banche dati con cadenza annuale. [S. Adams - "COMPARING THE IPC AND THE US CLASSIFICATION SYSTEMS FOR THE PATENT SEARCHER" - World Patent Information (2001), 23, p. 15 - 23]

documenti brevettuali US sono classificati in questo modo.

5. Differenze tra IPC7 e IPC8

Nel gennaio 2006 è entrata in vigore una nuova classificazione internazionale (denominata IPC8), il cui scopo è quello armonizzare le principali classificazioni, nonché di facilitare le ricerche dei documenti brevettuali.

Rispetto alle precedenti edizioni le principali novità riguardano:

- revisioni più frequenti¹⁰;
- la riclassificazione dei brevetti ogniqualvolta viene modificato un codice di classificazione;
- un maggior livello di dettaglio dei sottogruppi, in modo da armonizzare tutte le principali classificazioni (ECLA, F-Terms e USPC);
- la classificazione dei brevetti non solo sulla base delle rivendicazioni ma anche della descrizione;
- la creazione di due livelli: "core" e "advanced".

Nel livello "core" confluiscono anche i documenti classificati nel livello "advanced" ed è pertanto da utilizzare per effettuare ricerche complete. Il livello "advanced" (sottoposto a revisione ogni tre mesi) conterrà solo la documentazione minima PCT¹¹.

¹⁰ Ogni tre mesi per il livello "advanced" ed ogni tre anni per il livello "core"; le edizioni IPC 1-7 erano revisionate ogni 5 anni.

¹¹ La documentazione minima PCT comprende i brevetti dei seguenti uffici nazionali ed internazionali: Svizzera, Stati Uniti (pubblicati dopo il 1920), Francia, Germania e Regno Unito (pubblicati anche prima del 1920), Austria, Australia, Canada (pubblicati dopo il 1970), Giappone e Russia (solo riassunti in inglese, se disponibili), Olanda (pubblicati dal 1912), Belgio (pubblicati dal 1926) e Lussemburgo (pubblicati dal 1946), nonché tutte le domande EP, PCT ed i brevetti concessi EP. "GUIDELINES FOR EXAMINATION IN THE EUROPEAN PATENT OFFICE" - Cap. IX par. 2.1 (PCT minimum documentation) e 2.2 (Other national patent documents)

Le classificazioni IPC¹² 1 - 7 comprendono tutti i documenti brevettuali pubblicati dal 1968 al 2005 e quindi una ricerca completa, che comprenda anche quanto pubblicato prima del '68, è possibile solo con le classi IPC8 o ECLA.

Dal momento che ai documenti di recente pubblicazione non viene attribuito un codice ECLA (ma solo IPC), è consigliabile utilizzare l'IPC8 per avere la più ampia copertura¹³.

Non tutti gli Stati hanno adottato entrambi i livelli di classificazione dell'IPC8: alcuni attribuiscono ai brevetti nazionali solo il livello "core", pertanto una ricerca nel solo livello avanzato non può essere considerata completa.

6. La ricerca di brevetti giapponesi

Nel 1999 l'Ufficio Brevetti Giapponese (JPO) ha messo a disposizione gratuitamente sul web¹⁴ una libreria digitale sulla proprietà industriale (IPDL - "*Industrial Property Digital Library*"), nonché un sistema di ricerca e recupero dell'informazione brevettuale [9].

La ricerca può essere effettuata sia sui titoli ed i riassunti dei brevetti (mediante parole chiave e classificazione IPC), sia sul numero delle domande/brevetti concessi [10].

http://www.european-patent-office.org/legal/gui_lines/index.htm

A partire dal 1° aprile 2007 saranno inseriti anche i brevetti coreani.

¹² La settima edizione dell'IPC, entrata in vigore a partire dal gennaio 2000, era suddivisa in 8 sezioni, 20 sottosezioni, 120 classi, 628 sottoclassi e più di 69.000 gruppi (dei quali circa il 10% è costituito da gruppi principali, mentre il resto da sottogruppi) [E. Lo Conte - "LA CLASSIFICAZIONE INTERNAZIONALE DEI BREVETTI" - <http://www.netval.it/Documenti-PDF/loconte-1.pdf>]

¹³ Heiko Wongel - "IPC8 AND ECLA - COMPREHENSIVE SEARCH AND RETRIEVAL" - <http://academy.epo.org/schedule/2006/ic04/IC-04/IPC%20and%20ECLA/player.html>

¹⁴ Disponibile al sito http://www.ipdl.ncipi.go.jp/homepg_e.ipdl

Nel 2000 è stato aggiunto uno strumento automatico di traduzione dei brevetti (descrizione, rivendicazioni, didascalie dei disegni), anche se non preciso al 100%!

Inoltre, dal 2001 sono disponibili strumenti che consentono di effettuare ricerche sull'intera letteratura brevettuale giapponese mediante le classificazioni applicate dal JPO: *file index* (FI) e *F-terms*.

Il *file index* è la classificazione interna utilizzata dal JPO per classificare il contenuto tecnico dei documenti brevettuali e per consentire l'esecuzione delle ricerche sullo stato dell'arte in modo più efficiente [10].

La classificazione FI¹⁵ è applicata a tutti i brevetti e modelli d'utilità giapponesi ed è un'estensione della classificazione internazionale¹⁶.

La classificazione F-terms¹⁷, invece, non è assegnata a tutti i brevetti: la copertura dipende dal settore tecnologico.

Un'importante differenza tra IPC e F-terms riguarda proprio il metodo di classificazione dei documenti: da un singolo punto di vista tecnico secondo l'IPC, da molteplici punti di vista tecnici (materiale, operazione, scopo, ecc...) per F-terms. Il contenuto tecnico dei documenti brevettuali è analizzato dagli esaminatori e suddiviso in differenti codici tematici ("*theme code*"¹⁸ - rappresentanti un settore tecnologico); tali codici sono

¹⁵ La classificazione FI è costituita da una classe IPC, da un simbolo di sottodivisione (un numero di 3 cifre) e da un ulteriore simbolo (facoltativo), costituito da una lettera e denominato "*File discrimination symbol*".

¹⁶ La relazione tra FI e IPC può essere considerata simile a quella tra ECLA e IPC.

¹⁷ Gli F-terms sono termini tecnici utilizzati per migliorare l'efficienza delle ricerche di prior art e sono assegnati ad un brevetto non solo sulla base delle rivendicazioni ma anche della descrizione.

¹⁸ L'assegnazione degli F-terms è basata sulla classificazione FI. Ciascun codice tematico possiede una corrispondente copertura di codici FI.

suddivisi in codici di termine ("term code" - assegnati in funzione dei vari punti di vista tecnici).

Ciascun F-terms consiste di un codice tematico costituito da cinque cifre ed un codice di termine di quattro cifre.

Sul sito web IPDL è disponibile un servizio (denominato "Patent Map Guidance") in grado di fornire al ricercatore il supporto nella scelta della classificazione appropriata [12].

La maschera di ricerca "FI/F-terms Search"¹⁹ consente di effettuare ricerche nell'intera banca dati brevettuale giapponese. I campi ricercabili sono le classi FI, i codici tematici ed i codici di termine. È possibile combinare i criteri di ricerca mediante operatori booleani.

7. Conclusioni

Le ricerche brevettuali eseguite sulla classificazione ECLA non risultano del tutto complete, dal momento che non tutti i documenti sono classificati con questi codici e comunque sussiste un certo ritardo dalla pubblicazione della domanda di brevetto (che obbligatoriamente deve riportare la classificazione IPC) all'attribuzione di questi codici.

Il pregio della classificazione ECLA è di essere molto dettagliata e caratterizzata da una revisione dinamica.

Uno dei vantaggi dell'IPC8 è la copertura degli Stati, mentre uno svantaggio riguarda la sua applicazione da parte di differenti uffici nazionali, con conseguente vincolo all'interpretazione degli esaminatori di questi uffici.

Nelle ricerche di prior art non è da trascurare la cosiddetta letteratura non brevettuale (NPL).

¹⁹ In questa maschera di ricerca non è possibile effettuare alcuna ricerca per parole chiave.

Dai dati dell'UEB risulta che nel 2005 il 14,3% di tutte le citazioni²⁰ è riferito alla letteratura non brevettuale, con variazioni dall'1 al 60% (per le biotecnologie²¹), in funzione del settore tecnologico.

Bibliografia

[1] G. Conti et al. - "PROTEGGERE L'IDEA" - Franco Angeli Editore (2003), cap. 8 ("*Il brevetto come fonte d'informazione*"), p. 97 - 103

[2] "PATENT INFORMATION: BURIED TREASURE" - WIPO Magazine (2005), Jan./Feb., p. 8- 11

[3] "GUIDELINES FOR EXAMINATION IN THE EUROPEAN PATENT OFFICE" - Cap. III par. 2.1 (Completeness of the search)

<http://www.european-patent-office.org/legal/guidelines/index.htm>

[3] M. Stock, W. G. Stock - "INTELLECTUAL PROPERTY INFORMATION: A COMPARATIVE ANALYSIS OF MAIN INFORMATION PROVIDERS" - Journal of The American Society For Information Science and Technology (2006), 57 (13), p. 1794 - 1803

[4] S. van Dulken - "OBTAINING BUSINESS INFORMATION FROM PATENTS, DESIGNS AND TRADE MARKS" - Business Information Review (2006), 23 (2), p. 119 - 124

[5] D. Newton - "PATENTS INFORMATION: WHAT'S IN IT FOR THE BUSINESS INFORMATION USER?" - Business Information Review (1998), 15 (4), p. 248 - 253

[6] L. Falasco - "UNITED STATES PATENT CLASSIFICATION: SYSTEM ORGANIZATION" - World Patent Information (2002), 24, p. 111 - 117

²⁰ Da una ricerca è emerso che 2/3 delle citazioni brevettuali sono inserite dagli esaminatori; inoltre, il 40% dei brevetti possiede tutte citazioni aggiunte dagli esaminatori [J. Alcácer, M. Gittelman - "PATENT CITATIONS AS A MEASURE OF KNOWLEDGE FLOWS: THE INFLUENCE OF EXAMINER CITATIONS" - The review of Economics and Statistics (2006), 88 (4), p. 774 - 779]

²¹ I tre articoli più citati nel 2005 riguardano proprio le biotecnologie.

[7] L. Falasco - "BASES OF THE UNITED STATES PATENT CLASSIFICATION" - World Patent Information (2002), 24, p. 31 - 33

[8] I. Schellner - "SOURCES OF JAPANESE PATENT INFORMATION" - World Patent Information (2001), 23, p. 149 - 156

[9] T. Hayakawa - "THE INDUSTRIAL PROPERTY DIGITAL LIBRARY OF THE JAPANESE PATENT OFFICE" - World Patent Information (2000), 22, p. 35 - 37

[10] I. Schellner - "JAPANESE FILE INDEX CLASSIFICATION AND F-TERMS" - World Patent Information (2002), 24, p. 197 - 201

[11] "JAPANESE FI CLASSIFICATION AND F-TERMS NOW AVAILABLE IN ENGLISH", Epidos News 3/2001, October 2001, p. 7