

TEMI E ANALISI

Navigare nel mare di Scopus, Web of science e Google Scholar: L'avvio di una ricerca sulla vitalità delle discipline archivistiche e biblioteconomiche italiane

di Simona Turbanti

Impact is not the same as importance or significance
(Eugene Garfield, 1963)

1. Il contesto

Questo contributo nasce parallelamente all'avvio di una ricerca dal titolo *Internazionalizzazione e vitalità delle discipline biblioteconomiche italiane: per una ricerca quantitativa e qualitativa*, all'interno del 29. ciclo del dottorato in Scienze documentarie, linguistiche e letterarie - Curriculum in scienze librarie e documentarie - presso l'Università La Sapienza. Il lavoro, intrapreso da alcuni mesi, si propone di analizzare il livello di 'salute' degli studi del settore M-STO/o8, inteso come capacità di uscire fuori dalla propria nicchia ed essere presenti in aree disciplinari e/o linguistiche esterne alla propria.

Si rendono, innanzitutto, necessarie due precisazioni. Circa il titolo del progetto c'è da far presente che, benché inizialmente fosse stato ipotizzato un ambito limitato ai soli studi biblioteconomici, già nelle prime fasi della ricerca ci si è orientati sull'intera area di studi archivistica, bibliografica, biblioteconomica e di storia del libro compresa in M-STO/o8, come evidenziato nel titolo del presente contributo. Il secondo avvertimento riguarda l'avanzamento dello studio; quanto si cercherà di evidenziare in questa sede rappresenta solo l'esordio di una - anzi, più - analisi, che è stato preceduto, come naturale, da una mappatura dello stato dell'arte a livello italiano e internazionale, per alcuni aspetti ancora in corso. Pertanto, i risultati esposti in questa prima rassegna hanno caratteristiche provvisorie e potranno essere illustrati e contestualizzati correttamente solo al termine della ricerca stessa.

SIMONA TURBANTI, Università di Pisa, Sistema bibliotecario di ateneo - Dottoranda presso il Dipartimento di Scienze documentarie, linguistico-filologiche e geografiche, Sapienza Università di Roma, viale Regina Elena 295, 00161 Roma, email simona.turbanti@uniroma1.it.

Ultima consultazione siti web: 31 luglio 2014



2. La 'citation culture'

Senza voler ripercorrere il percorso che, dalla valutazione paritaria della ricerca ha condotto nel tempo alla bibliometria, riassunto efficacemente nel precedente numero di questa rivista¹, è opportuno però soffermarsi brevemente sul significato di quella «citation culture», espressione coniata da Paul Wouters² che afferma

This subculture has unwittingly and subtly changed core concepts of modern science such as scientific quality and influence. Because of the citation culture, *being cited* has profoundly changed its meaning over the last two decades, with a number of consequences for scientists³.

Spingendosi oltre, l'autore sostiene l'esistenza di più di una «citing culture», diversa a seconda del tipo di comunità di scienziati presso cui è ospitata: i matematici tendono a non citare molte pubblicazioni, mentre gli studiosi di discipline biomediche abbondano con i riferimenti, così come gli storici e i ricercatori di letteratura, pur con modalità differenti gli uni dagli altri. Anche nelle varie riviste scientifiche cambiano le «citing culture», non solo per lo stile tipografico, ma riguardo al tipo di citazione, al suo numero, alla posizione nel testo⁴. A questo proposito, anche Gustav Nelhans parla di tre finalità diverse della citazione a seconda dei settori disciplinari: mettere in risalto l'influenza - di qualcuno - nelle scienze naturali, sottolineare aspetti testuali all'interno del settore umanistico (fino a divenire spesso il *subject* della ricerca), infine, rafforzare le argomentazioni nel campo delle scienze sociali⁵.

Wouters sottolinea inoltre come il graduale sviluppo dell'abitudine a citare, negli scritti, i lavori propri o di altri studiosi abbia portato inevitabilmente all'utilizzo di questi «citation data» e, di conseguenza, a valutare un lavoro in rapporto al numero di citazioni ottenute. Nell'introduzione della sua tesi di dottorato lo studioso olandese avverte che

the citation culture is not a simple aggregate or derivative of citing culture in science. The citation as used in scientometric analysis and science and technology indicators is not identical to the reference produced at the scientist's desk. This is the first claim of my study: the citation is the product of the citation indexer, not of the scientist. Citation analysis has only been feasible on a discernable scale since the invention of computerized citation indexes⁶.

1 Cfr. Luca Lanzillo, *Bibliotecari, bibliometria e valutazione della ricerca: riscoprire una competenza per valorizzare una professione*, «AIB studi», 54 (2014), n. 1, p. 51-60, <<http://aibstudi.aib.it/article/view/9497>>, DOI: 10.2426/aibstudi-9497.

2 Paul Wouters, *The citation culture*, tesi di dottorato, Università di Amsterdam, 1999, <<http://dare.uva.nl/document/487448>>.

3 *Ivi*, p. 2.

4 *Ivi*, p. 3.

5 Gustav Nelhans, *Qualitative scientometrics*, Proceedings of the IATUL conferences, Paper 6, 2014, <<http://docs.lib.purdue.edu/iatul/2014/plenaries/6>>.

6 P. Wouters, *The citation culture cit.*, p. 4.

Quanto descritto da Wouters e da altri⁷ è l'evoluzione del significato della citazione nel corso del tempo. Durante il secolo scorso, man mano che la mole della produzione scientifica aumenta, la citazione si prefigura come una naturale operazione, divenendo «la moneta corrente nel commercio della comunicazione scientifica ufficiale. Moneta di piccolo taglio (costa poco citare), ma dal potere d'acquisto simbolico non indifferente»⁸.

Ma è solo agli inizi degli anni '60, con la nascita dello Science Citation Index ad opera di Eugene Garfield, che vengono svelate al mondo esterno a quello degli studiosi le potenzialità racchiuse nelle citazioni; da quel momento in avanti, infatti, «la citazione ha cominciato a vivere una vita propria sganciata da quella dei documenti che costituivano la sua naturale dimora»⁹ rappresentando il mezzo attraverso il quale sondare un reticolo, potenzialmente infinito, di richiami e rinvii. Da lì a diventare la componente principale della 'pratica della misura' del peso scientifico di riviste e - pur non del tutto propriamente - di studiosi e ricercatori il passo è stato breve. Seguendo questa evoluzione della citazione, nella portata e nel significato, si capisce il motivo per cui Wouters definisce la «citation culture» un ibrido di scienza e politica.

3. Tre 'registri' noti sul palcoscenico bibliometrico: Scopus, Web of science e Google Scholar

Come già accennato, il *focus* della ricerca consiste nell'analisi del livello di internazionalizzazione e di vitalità delle discipline archivistiche e biblioteconomiche italiane attraverso indagini effettuate sia con indicatori bibliometrici quantitativi sia mediante metodi valutativi di tipo qualitativo. Il primo passo compiuto è stato l'interrogazione delle due principali banche dati citazionali interdisciplinari esistenti e utilizzate a livello internazionale: Web of science¹⁰ - da ora in avanti WoS - di Thomson Reuters e Scopus¹¹ di Elsevier. Per ricordare, assai sinteticamente, le principali caratteristiche dei due strumenti basti dire che WoS ricerca in oltre 12.300 riviste, circa 150.000 atti di convegni e 30.000 pubblicazioni monografiche, con una netta prevalenza di titoli dell'area nord americana, mentre il più 'giovane' Scopus copre circa 19.500 titoli di periodici dei quali molti di ambito europeo, oltre 350 collezioni di monografie (e, a partire dal 2013, anche monografie singole), quasi 5.000.000 di *conference paper* e brevetti. Come noto, entrambe le fonti condividono un elemento di forte criticità per il settore umanistico e delle *social sciences*, ossia la forte predominanza di riviste occidentali in lingua inglese (caratteristica più accentuata in WoS) e la carenza di strumenti/filtri di disambiguazione degli autori omonimi. Nonostante tutti i limiti, abbiamo ritenuto l'interrogazione di queste due banche dati un punto di partenza irrinunciabile per il nostro studio.

Il terzo 'registra' coinvolto è Google Scholar¹², il motore di ricerca dedicato ai risultati della produzione scientifica accessibile gratuitamente in rete, lanciato quasi con-

7 Si veda Nicola De Bellis, *La citazione bibliografica nell'epoca della sua riproducibilità tecnica: bibliometria e analisi delle citazioni dallo Science Citation Index alla Cybermetrica*, [ultima revisione 31/05/2005], <<http://www.bibliotecheoggi.it/content/CITAZIONE.pdf>>.

8 *Ivi*, p. 9.

9 *Ivi*, p. 12.

10 <<http://wokinfo.com/>>.

11 <<http://www.elsevier.com/online-tools/scopus>>.

12 <<http://scholar.google.it/>>.

temporaneamente a Scopus. Non trattandosi di una base dati bibliografica, non presenta limiti nel tipo di fonti consultate, né nella loro quantità e qualità; viene indicizzata, infatti, una vasta letteratura scientifica online, in costante crescita, di cui sono ignoti i dettagli circa la copertura e le fonti dell'analisi citazionale. I problemi legati alla mancanza di controllo di autorità per i nomi degli autori, riscontrabili anche in WoS e nel database di Elsevier, diventano pressanti nel mondo Google. Su questo aspetto della ricerca ci soffermeremo più avanti. Da ricordare, inoltre, l'esistenza di Google Scholar Citations, «a simple way for authors to keep track of citations to their articles»¹³, di cui avremo modo di parlare durante l'esposizione.

Una dettagliata analisi, pur ferma al 2005, dei tre strumenti si deve a Péter Jacsó che ha effettuato numerosi test di confronto mediante la ricerca di documenti che citano: Garfield; un articolo dello stesso pubblicato in *Science* nel 1955; la rivista *Current Science*; i 30 articoli più citati di *Current Science*. Le conclusioni cui è giunto Jacsó hanno dimostrato l'esistenza di molti dati presenti esclusivamente in una delle tre fonti, la maggior parte dei quali rilevanti ai fini della ricerca¹⁴.

Scopus, WoS e Google Scholar sono stati oggetto anche degli studi di Lokman I. Meho e Kiduk Yang¹⁵. I due studiosi hanno ricercato le citazioni di circa 1.300 lavori scientifici pubblicati da un gruppo di docenti dell'area LIS della School of Library and information science dell'Indiana University, quantificando il tempo necessario per la raccolta e l'esame di oltre 10.000 documenti citanti, o presumibilmente citanti (100 ore di lavoro in WoS, 200 in Scopus e 3.000 in Google Scholar). L'analisi comparativa compiuta da Meho e Yang ha evidenziato l'opportunità di affiancare a WoS la banca dati di Elsevier, in grado di influenzare in modo significativo il *ranking* e il numero delle citazioni degli studiosi presi in esame. Di Scholar viene sottolineata la buona copertura per quanto riguarda gli atti di convegni, in gran parte grazie all'auto-archiviazione online dei *paper*, e le riviste internazionali in lingua non inglese. Secondo i due autori, però, «Google Scholar is also very cumbersome to use and needs significant improvement in the way it displays search results and the downloading capabilities it offers for it to become a useful tool for large-scale citation analyses»¹⁶ e, a differenza di Scopus, ha scarso peso sul *ranking* dei soggetti analizzati. Nonostante questi limiti - che, giova ricordare, si riferiscono a una versione non recente dello strumento - Scholar si rivela indispensabile per mostrare l'impatto internazionale degli studiosi ed il suo utilizzo insieme a WoS e Scopus permette di avere un quadro più completo dei rapporti scientifici tra una disciplina (la LIS) ed altre (informatica, ingegneria, scienze cognitive, ecc.).

Infine, un'indicazione metodologica importante che emerge dal lavoro di Meho e Yang è la necessità, per chi conduce indagini simili specie su grandi numeri e in presenza di autori omonimi, di avere a disposizione la lista completa delle pubblicazioni degli studiosi analizzati.

13 Cfr. <<http://scholar.google.com/intl/en/scholar/citations.html>>.

14 Péter Jacsó, *As we may search: comparison of major features of the Web of science, Scopus and Google Scholar citation-based and citation-enhanced databases*, «Current science», 89 (2005), n. 9, 10, p. 1537-1547, <<http://choo.fis.utoronto.ca/FIS/courses/LIS1325/Readings/jacso.pdf>>.

15 Lokman I. Meho; Kiduk Yang, *Impact of data sources on citation counts and rankings of LIS faculty: Web of science versus Scopus and Google scholar*, «Journal of the American Society for Information Science and Technology», 58 (2007), n. 13, p. 2105-2125, DOI: 10.1002/asi.20677.

16 *Ivi*, p. 2123.

4. La ricerca nei tre strumenti: metodo seguito, confronti di uso, limiti riscontrati

Il campione analizzato nella prima fase del lavoro è costituito dai docenti e ricercatori del settore M-STO/o8, i cui nominativi sono stati scaricati dal sito del Ministero dell'istruzione, dell'università e della ricerca (77 nominativi, di cui 17 ordinari, 3 straordinari, 24 associati, 33 ricercatori inclusi i ricercatori a tempo determinato)¹⁷. Per tentare di avere un quadro più rappresentativo del panorama scientifico nazionale, tale elenco sarà in seguito integrato con un certo numero di professionisti, non docenti, del settore bibliotecario, selezionati in base alla quantità di pubblicazioni prodotte e alla loro continuità nel tempo.

Come primo passo si è provveduto ad effettuare la ricerca di lavori - articoli di riviste, contributi monografici, interventi in conferenze, recensioni a pubblicazioni altrui - dei 77 studiosi indicizzati all'interno di WoS e Scopus. In WoS l'interrogazione è stata condotta dalla maschera *Basic search*, selezionando *Author* nel menu a tendina laterale; in caso di cognomi di uso molto comune i numerosi risultati ottenuti sono stati ristretti in base alle *Web of science categories* (filtro che appare, insieme ad altri criteri di affinamento, nella colonna sinistra a fianco dei documenti). Analogamente, nel database di Elsevier si è partiti dalla maschera *Author search* senza applicare alcun filtro; anche in questo caso, un numero troppo alto di *record* è stato raffinato mediante la *Subject area*. In nessuno dei due strumenti ci si è affidati all'opzione della ricerca per affiliazione, dato spesso assente o errato. Le ricerche sono state condotte digitando il cognome seguito dall'iniziale del nome - scelta obbligata se si opta per la *Author search* di WoS - così da evitare di perdere nel 'mare' dei dati *record* in cui il nome dell'autore è registrato con la sola iniziale. In presenza di cognomi formati da più elementi occorre ripetere la ricerca utilizzando tutte le forme.

A questo proposito, prima di proseguire nella descrizione della ricerca vale la pena spendere qualche parola su un'assenza che pesa in modo rilevante nei due database citazionali (e, come vedremo in seguito, ancor di più in Google Scholar), il controllo di autorità. Come appare evidente a chiunque abbia provato, almeno una volta, a effettuare una ricerca in WoS e Scopus, il riconoscimento degli autori non è sempre agevole. Il mancato raggruppamento delle forme varianti di un nome in una scheda di autorità, con il conseguente rinvio da tutte le varianti alla forma del nome accettata, rende spesso incerta l'identificazione dell'autore che si sta cercando. Se si fatica a non perdere l'orientamento all'interno di cataloghi di biblioteca privi di *authority lists*, diventa assai difficile, a maggior ragione, muoversi in contenitori dai confini in continua espansione e non sottoposti alla mediazione del processo catalografico come i database internazionali¹⁸.

Benché si sia lontani in entrambi i casi dall'aver informazioni corrette sugli autori, la situazione di Scopus e WoS non è però identica; nella base dati di Elsevier il livello delle informazioni è più affidabile grazie all'assegnazione, ove possibile, di un *Author Identifier*¹⁹ (un numero univoco, una sorta di ISADN, ottenuto con un algoritmo) e alla collaborazione degli autori stessi, invitati a fornire *feedback* e segnalare errori. In WoS, invece, si spingono gli autori - o le istituzioni presso le quali sono in servizio - a provvedere autonomamente alla propria, corretta, identificazione, mediante *ResearcherId*²⁰,

17 Cfr. <<http://cercauniversita.cineca.it/php5/docenti/cerca.php>>.

18 Si veda Simona Turbanti, *Cui prodest libraries authority work?*, «JLIS», 5 (2014), n. 2, p. 55, <<http://leo.cineca.it/index.php/jlis/article/view/10062>>, DOI: 10.4403/jlis.it-10062.

19 Cfr. <http://help.scopus.com/Content/h_autsrch_intro.htm>.

20 Cfr. <<http://wokinfo.com/researcherid/>>.

una comunità online dove gli autori si creano profili personalizzati e interagiscono tra loro; è prevista, inoltre, la possibilità di un'integrazione tra RID e l'iniziativa ORCID²¹.

Un altro, grave, limite delle due basi dati citazionali è rappresentato dalla copertura delle riviste indicizzate. Come noto, la copertura reale di una rivista non è sempre coincidente con i dati ufficialmente dichiarati, nei cataloghi così come nelle banche dati; capita non di rado ci siano lacune nelle annate non rese esplicite. All'interno di un database citazionale come Scopus e WoS questo provoca, ovviamente, un possibile disallineamento dei risultati ottenuti interrogando le due banche dati per gli stessi autori; i lavori di uno studioso possono, infatti, non solo essere presenti in una e non nell'altra risorsa, a seconda delle testate indicizzate, ma anche risultare assenti, nonostante la rivista sia spogliata, a causa di lacune nel posseduto=indicizzato. Questo tipo di 'buco' non sempre facilmente appurabile, rischia quindi di falsare l'indice di impatto citazionale degli autori; se poi si somma l'incidenza di eventuali lacune all'ancora scarsa presenza di testate italiane di area socio-umanistica nelle due basi dati, il quadro diventa poco incoraggiante.

Per rendersi conto del tipo di riviste italiane indicizzate nei due strumenti, nelle quali è stato pubblicato almeno un lavoro di docenti e ricercatori di M-STO/08, può essere utile fornirne un elenco.

Riviste presenti in Scopus:

Archivio storico italiano (Olschki);
History of education and children's literature (Edizioni Università di Macerata);
Passato e presente (Angeli);
Rinascimento (Olschki);
Studi storici (Carocci);
L'universo (Istituto geografico militare).

Riviste presenti in WoS:

Archivio storico italiano (Olschki);
Aevum (Vita e pensiero);
Critica letteraria (Loffredo);
Historica (Amministrazione Comunale di Reggio Calabria);
History of education and children's literature (Edizioni Università di Macerata);
Nuova rivista storica (Società editrice Dante Alighieri);
Il ponte (Il ponte);
Quaderni medievali (Dedalo);
Quaderni storici (Il Mulino);
Rassegna storica del Risorgimento (Istituto per la storia del Risorgimento italiano);
Rassegna storica toscana (Olschki);
Rinascimento (Olschki);
Rivista di letteratura italiana (Serra);
Rivista di letterature moderne e comparate (Pacini);
Rivista storica italiana (Edizioni scientifiche italiane);
Studi e problemi di critica testuale (Serra);
Studi medievali (Centro italiano di studi sull'alto Medioevo);
Studi piemontesi (Centro di studi piemontesi);
Studi storici (Carocci).

²¹ Cfr. <<http://orcid.org/>>.

I titoli presenti in entrambe le banche dati sono dunque solo 4 (*Archivio storico italiano, History of education and children's literature, Rinascimento, Studi storici*) e, in alcuni casi, si è riscontrata una copertura delle annate non coincidente. Come appare evidente, si tratta di periodici di area storica e letteraria, al cui interno trovano spazio contributi di storia del libro e della stampa; totalmente assenti risultano le riviste di area archivistica e biblioteconomica²².

Come accennato più volte, la ricerca è stata compiuta anche in Google Scholar e Google Scholar Citations (integrati, ove necessario, da Google e Google Books). Nonostante siano stati, per il momento, esaminati in modo completo solo alcuni autori tra i 77 totali, il quadro risultante ha rivelato notevoli complessità che hanno costretto a un ripensamento di alcune scelte e portato a una chiara definizione di punti fermi per 'navigare' senza perdersi nella mole di informazioni.

La prima difficoltà nell'uso del celebre motore di ricerca specializzato è rappresentata dall'ambiguità dei risultati che rende faticosa la lettura dei dati, specie per autori molto prolifici. Il numero di lavori di un autore risultante da una ricerca lanciata in Google Scholar e Google Scholar Citations non corrisponde, necessariamente, al numero di entità - per dirla in termini catalografici - realmente esistenti e attribuibili a quell'autore. Oltre al fenomeno dei doppioni di una medesima opera vanno infatti considerate le citazioni di una medesima opera duplicate (ad esempio, una citazione pubblicata in *preprint* e all'interno di una rivista, conteggiata due volte) e gli errori nelle informazioni bibliografiche delle pubblicazioni. Per questo motivo, nella fase di analisi dei risultati di un autore si rende spesso necessario un controllo delle informazioni negli OPAC e/o in Google Books e Google. Altro fattore di criticità è rappresentato dall'alto tasso di errori e duplicazioni nelle forme dei nomi degli autori, su cui abbiamo già avuto modo di riflettere parlando di WoS e Scopus; se il controllo dei dati è già un'impresa ardua all'interno di contesti vasti ma delimitati, diviene inevitabilmente impossibile nel mondo senza confini della rete. Un rimedio, però, esiste ed è la già citata applicazione Citations che consente a ogni utente di creare un proprio profilo autore, con i dati istituzionali, una foto, il numero e gli indici di citazioni (Indice H e i10-index). L'unica condizione richiesta per l'attivazione del profilo, che si può decidere di rendere pubblico o, al contrario, di tenere privato, è il possesso di un indirizzo e-mail valido su cui viene effettuata da Google una verifica in automatico. Qualora l'utilizzo di Google Citations diventasse una pratica comune per tutti gli autori - includendo in tale categoria i responsabili di qualsiasi tipo di creazione intellettuale in rete, anche i documenti Word e presentazioni in Power Point - la maggior parte dei problemi di *authority control* sarebbe superata grazie all'identificazione delle opere alla fonte.

Da tenere infine presente che in Citations, in fondo ad ogni pagina utente, viene specificato che «date e numero di citazioni sono delle stime e sono determinati automaticamente da un programma software».

5. La ricerca: i primi risultati

Illustrati i metodi di ricerca e alcuni tra i principali problemi riscontrati nell'utilizzo delle fonti - abbiamo omesso, tra gli altri, la complessa questione delle citazioni non incluse nell'indicizzazione e quelle indicizzate ma non riconosciute come tali

²² Recentissima la notizia dell'inclusione di AIB studi in *Scopus*, cfr. editoriale di Giovanni Di Domenico e Gabriele Mazzitelli, *AIB Studi, le biblioteche e la valutazione della ricerca*, «AIB studi», 54 (2014), n. 1, p. 7-9, <<http://aibstudi.aib.it/article/view/10022/9406>>, DOI: 10.2426/aibstudi-10022.

che richiedono, specie in Google Scholar, ricerche aggiuntive – torniamo all’oggetto della ricerca.

Per ognuno dei docenti e ricercatori del settore M-STO/o8 sono stati raccolti e organizzati in una tabella Excel i seguenti dati:

- numero di lavori scientifici segnalati in WoS, Scopus e Google Scholar: per le due banche dati è stato fornito un numero totale di pubblicazioni, segnalando tra parentesi il numero delle eventuali recensioni edite di altri lavori; in Scholar, invece, si è proceduto a scorporare la somma totale in 4 categorie (articoli, libri, parti di libri/atti di convegni, contributi vari sul web; all’interno degli articoli sono state specificate le recensioni);
- numero complessivo di citazioni ricevute dai lavori;
- arco cronologico di riferimento entro il quale si collocano i lavori;
- titoli delle riviste in cui sono pubblicati i lavori;
- lingua di pubblicazione dei lavori.

Questo il sunto dei risultati, aggiornati al mese di giugno scorso.

In Scopus risultano presenti 42 lavori, costituiti per la maggior parte da articoli o recensioni in riviste internazionali di archivistica e biblioteconomia in lingua inglese, 2 in lingua francese e alcune testate italiane di ambito storico. In WoS compaiono invece 300 lavori, anche in questo caso composti in maggioranza da articoli in riviste di archivistica e biblioteconomia internazionali; maggiore anche il tasso di recensioni all’interno dei lavori rispetto alla banca dati di Elsevier (205 sul totale, circa il 68%, a fronte di 5 recensioni sui 42 lavori presenti in Scopus che corrisponde a circa il 12%).

Il totale delle citazioni ricevute dai 42 lavori di docenti/ricercatori in Scopus è pari a 32, mentre la quota sale a 66 sui quasi 300 contributi presenti in WoS; con tutta probabilità il rapporto tra numero di lavori e numero di citazioni è più basso in quest’ultima banca dati a causa della forte presenza di recensioni, solitamente scarsamente citate. Le 32 citazioni in Scopus sono così ripartite tra i 77 docenti/ricercatori del settore: 18 di lavori di ordinari, 4 di associati e 10 citazioni di ricercatori; anche le 66 citazioni in WoS vedono una predominanza del numero di citazioni di contributi di ordinari (44), mentre il numero scende tra i ricercatori (16) e ancora di più fra gli associati (6).

Come prevedibile, si tratta di risultati quantitativamente poco rilevanti; in settori disciplinari specialistici come il nostro, con scarsa copertura nelle banche dati citazionali, pochi *core journals* e un numero ridotto di pubblicazioni monografiche indicizzate, è infatti assai difficile incidere con una presenza forte. Con tali presupposti il dato numerico non potrà essere significativo a fini valutativi se sganciato da un attento ragionamento sulle fonti (e, come vedremo nelle conclusioni, è in questo aspetto che risiede, a nostro parere, l’interesse di rilevazioni citazionali di questo tipo).

Passiamo ora a Google Scholar. Premesso che la ricerca è tuttora in corso e che non si è in grado, se non con un’analisi approfondita, di valutare a fondo la pertinenza dei risultati ottenuti, abbiamo deciso di muoverci seguendo alcuni criteri in grado di semplificare, almeno in parte, la raccolta e interpretazione delle informazioni ottenute (*record* o *master record*):

- analisi delle pagine di autori in Google Scholar Citations, quando esistenti. Da tale applicazione, aggiornata automaticamente, è, infatti, possibile visualizzare i record provenienti da Scholar in modo più pulito e facilmente leggibile; qualora non sia disponibile la pagina in Citations si è proceduto lanciando una ricerca a partire da Scholar, digitando nome - sia in forma sciolta che puntata - cognome dell’autore racchiuso tra virgolette;

- esportazione dei *record* ottenuti in EndNote web²³, un programma per la gestione dei riferimenti bibliografici;
- eliminazione dei *record* che descrivono medesime opere o contributi (duplicati);
- eliminazione dei *record* che descrivono parti di opere o contributi già descritti unitariamente e somma del numero di citazioni di ciascun record;
- fusione dei *record* in cui sono descritte nuove edizioni e/o ristampe di una medesima opera o contributo e somma del numero delle citazioni di ciascun *record*;
- esclusione delle citazioni di volumi in cui un autore ha scritto un articolo/parte e di volumi in cui un autore ha redatto un'introduzione, prefazione/postfazione (sono contate, invece, le citazioni di volumi con curatela);
- esclusione delle citazioni di tesi di qualsiasi grado di cui l'autore è relatore/*tutor*.

I docenti/ricercatori analizzati a partire da Google Scholar Citations sono e saranno contrassegnati da un asterisco nella tabella Excel, così da distinguerli da coloro che sono privi della pagina utente.

Nonostante il tentativo di razionalizzazione, il percorso rimane in salita; citiamo, come esempio, il caso di un docente ordinario del settore, dalla vasta produzione scientifica, privo di profilo su Citations e analizzato, dunque, in Scholar. In risposta alla nostra interrogazione vengono forniti circa 550 risultati; se i *record* delle pagine iniziali sono interpretabili abbastanza agevolmente (ma dovendo ricorrere spesso all'aiuto di OPAC, *institutional repositories*, *open archives*, siti di editori, Google Books, Google per la corretta interpretazione, controllo e pulizia dei dati), dopo la decima pagina le informazioni divengono di difficile comprensione. Si mescolano, infatti, notizie relative a opere dell'autore, spesso duplicati di altre già contenute nella pagine precedenti, notizie relative a citazioni, notizie tratte da pagine web di diverso tipo in cui viene nominato l'autore, citazioni ricavate dall'interno di libri o fascicoli di riviste a cui non è possibile accedere e citazioni, infine, che sembrano non rimandare ad alcun contributo legato all'autore.

Un altro caso esemplificativo è fornito da un ricercatore di M-STO/o8 cui vengono assegnate, in Google Scholar Citations, una ventina di pubblicazioni di responsabilità di un omonimo, la cui identità è stata appurata grazie alla voce d'autorità presente in SBN. Questa erronea attribuzione di opere di un autore diverso, peraltro molto citato, porta inevitabilmente un aumento del numero di citazioni e degli indici bibliometrici del ricercatore M-STO/o8 e l'alterazione del relativo quadro. Naturalmente, questo tipo di errori potrebbe essere ridotto, se non addirittura eliminato, dall'intervento diretto degli autori sui dati dei propri profili Citations, iniziativa che sarebbe auspicabile venisse portata avanti anche su input delle istituzioni di ricerca.

6. Conclusioni

Questa prima ricognizione non fa che confermare quanto dimostrato dagli studi condotti da Meho e Yang²⁴, vale a dire il bisogno di ricorrere a tutte e tre le risorse - WoS, Scopus, Scholar - per ottenere un quadro più completo possibile della produzione scientifica di uno studioso e della relative citazioni. Risulta inoltre evidente la complessità di uso degli strumenti citazionali e, soprattutto, la difficoltà di ricavarne risultati quantitativi rilevanti. Come già accennato, infatti, in contesti disciplinari poco rappresentati nel panorama citazionale come il nostro, il dato numerico

²³ <<http://endnote.com/>>.

²⁴ L.I. Meho; K. Yang, *Impact of data sources on citation counts and rankings of LIS faculty: Web of science versus Scopus and Google scholar cit.*

non potrà essere significativo a fini valutativi né l'analisi del risultato avere una qualche utilità, se avulsa dall'analisi dettagliata delle fonti. La priorità emersa chiaramente al termine di questa prima fase della ricerca riguarda, dunque, la collocazione dei dati quantitativi all'interno di un corretto quadro di riferimento; solo in tal modo si potrà, infatti, cercare di rappresentare i fenomeni di presenza e impatto - o, in altre parole, di internazionalizzazione e vitalità delle discipline LIS italiane - nel modo più 'oggettivo' possibile.

Siamo convinti, peraltro, che proprio quest'ultimo aspetto rivestirà un ruolo importante in relazione alle politiche editoriali e alle strategie connesse, sia a livello di intere aree scientifiche sia a livello individuale dei singoli studiosi.

RIFERIMENTI BIBLIOGRAFICI

- [1] Baccini Alberto, *Valutare la ricerca scientifica: uso e abuso degli indicatori bibliometrici*. Bologna: Il Mulino, 2010.
- [2] Bakalbas Nisa; Bauer Kathleen; Glover Janis; Wang Lei, *Three options for citation tracking: Google Scholar, Scopus and Web of Science*, «Biomedical digital libraries», vol. 3 (2006), n. 7, <<http://www.bio-diglib.com/content/3/1/7>>, DOI:10.1186/1742-5581-3-7.
- [3] Bosman Jeroen; Mourik Ineke van; Rasch Menno; Sieverts Eric; Verhoeff Huib, *Scopus reviewed and compared: the coverage and functionality of the citation database Scopus, including comparisons with Web of Science and Google Scholar*, report, Utrecht University Library (2006), <<http://dspace.library.uu.nl/handle/1874/18247>>.
- [4] Capaccioni Andrea; Spina Giovanna, *Italian HSS journals in Journal Citation Reports (JCR) and in SCImago Journal Rank (SJR): data and first analysis*, «JLIS», 3 (2012), n. 1, <<http://leo.cineca.it/index.php/jlis/article/view/4787>>, DOI: 10.4403/jlis.it-4787.
- [5] Cassella Maria; Bozzarelli Oriana, *Nuovi scenari per la valutazione della ricerca tra indicatori bibliometrici citazionali e metriche alternative nel contesto digitale*, «Biblioteche oggi», 29 (2011), n. 2, p. 66-78, <<http://www.bibliotecheoggi.it/content/n20110206601.pdf>>.
- [6] De Bellis Nicola, *La citazione bibliografica nell'epoca della sua riproducibilità tecnica: bibliometria e analisi delle citazioni dallo Science Citation Index alla Cybermetrica*, [ultima revisione 31/05/2005], <<http://www.bibliotecheoggi.it/content/CITAZIONE.pdf>>.
- [7] De Bellis Nicola, *Introduzione alla bibliometria: dalla teoria alla pratica*. Roma: Associazione italiana biblioteche, 2014.
- [8] De Robbio Antonella, *Analisi citazionale e indicatori bibliometrici nel modello Open Access*, «Bollettino AIB», 47 (2007), n. 3, p. 257-287, <<http://bollettino.aib.it/article/view/5246>>, NBN: urn:nbn:it:aib-4777.
- [9] Di Domenico Giovanni; Mazzitelli Gabriele, *AIB Studi, le biblioteche e la valutazione della ricerca*, «AIB studi», 54 (2014), n. 1, p. 7-9, <<http://aibstudi.aib.it/article/view/10022/9406>>, DOI: 10.2426/aibstudi-10022.
- [10] Ebrahim Nader Ale [et al.], *Equality of Google Scholar with Web of Science Citations; case of Malaysian engineering highly cited papers*, «Modern applied science», 8 (2014), n. 5, p. 63-69, <<http://www.ccsenet.org/journal/index.php/mas/article/view/37570>>, DOI: 10.5539/mas.v8n5p63.

- [11] Faggiolani Chiara; Solimine Giovanni, *La valutazione della ricerca, la bibliometria e l' albero di Bertoldo*, «AIB studi», 52 (2012), n. 1, p. 57-63, <<http://aibstudi.aib.it/article/view/6290>>, DOI: 10.2426/aibstudi-6290.
- [12] Galimberti Paola, *Qualità e quantità: stato dell' arte della valutazione della ricerca nelle scienze umane in Italia*, «JLIS», 3 (2012), n. 1, <<http://leo.cineca.it/index.php/jlis/article/view/5617>>, DOI: 10.4403/jlis.it-5617.
- [13] Jacsó Péter, *As we may search: comparison of major features of the Web of science, Scopus and Google Scholar citation-based and citation-enhanced databases*, «Current science», 89 (2005), n. 9, 10, p. 1537-1547, <<http://choo.fis.utoronto.ca/FIS/courses/LIS1325/Readings/jacso.pdf>>.
- [14] Jacsó Péter, *Google Scholar revisited*, «Online information review», 32 (2008), n. 1, p. 102-114, <<http://www.emeraldinsight.com/doi/full/10.1108/14684520810866010>>, DOI: 10.1108/14684520810866010.
- [15] Kear Robin; Colbert-Lewis Danielle, *Citation searching and bibliometric measures: resources for ranking and tracking*, «College & research libraries news», 72 (2011), n. 8, p. 470-474, <<http://crln.acrl.org/content/72/8/470.full>>.
- [16] Kousha Kayvan; Thelwall Mike; Rezaie Somayeh, *Assessing the citation impact of books: the role of Google Books, Google Scholar, and Scopus*, «Journal of the American Society for information science and technology», 62 (2011), n. 11, p. 2147-2164, <<http://onlinelibrary.wiley.com/doi/1002/asi.21608/full>>, DOI: 10.1002/asi.21608.
- [17] La Guardia Cheryl, *E-Views and reviews: Scopus vs Web of Science*, «Library journal», 15 January 2005, <<http://lj.libraryjournal.com/2005/01/ljarchives/e-views-and-reviews-scopus-vs-web-of-science/>>.
- [18] Lanzillo Luca, *Bibliotecari, bibliometria e valutazione della ricerca: riscoprire una competenza per valorizzare una professione*, «AIB studi», 54 (2014), n. 1, p. 51-60, <<http://aibstudi.aib.it/article/view/9497>>, DOI 10.2426/aibstudi-9497.
- [19] Meho Lokman I.; Sugimoto Cassidy R., *Assessing the scholarly impact of information studies: a tale of two citation databases—Scopus and Web of Science*, «Journal of the American Society for information science and technology», 60 (2009), n. 12, p. 2499-2508, <<http://onlinelibrary.wiley.com/doi/1002/asi.21165/full>>, DOI: 10.1002/asi.21165.
- [20] Meho Lokman I.; Sugimoto Cassidy R., *Mapping the intellectual impact of library and information science research through citations: a tale of two database: Scopus and Web of science*, «Proceedings of the American Society for information science and technology», 44 (2007), n. 1, p. 1-7, <<http://onlinelibrary.wiley.com/doi/1002/meet.1450440276/full>>, DOI: 10.1002/meet.1450440276.
- [21] Meho Lokman I.; Yang Kiduk, *Impact of data sources on citation counts and rankings of LIS faculty: Web of science versus Scopus and Google scholar*, «Journal of the American Society for information science and technology», 58 (2007), n. 13, p. 2105-2125, <<http://onlinelibrary.wiley.com/doi/1002/asi.20677/full>>, DOI: 10.1002/asi.20677.
- [22] Mayr Philipp, *Google scholar: how deep does this search engine dig?*, conference poster in *International Conference of the International Society for Scientometrics and Informetrics, 10th, Stockholm (Sweden), 24-28 July 2005*, <<http://hdl.handle.net/10760/6649>>.
- [23] Nelhans Gustav, *Qualitative scientometrics*, Proceedings of the IATUL conferences, Paper 6, 2014, <<http://docs.lib.purdue.edu/iatul/2014/plenaries/6>>.
- [24] Piazzini Tessa, *Gli indicatori bibliometrici*, «JLIS», 1 (2010), n. 1, p. 63-86, <<http://leo.cineca.it/index.php/jlis/article/view/24>>, DOI: 10.4403/jlis.it-24.

- [25] Salini Silvia, *La valutazione della ricerca scientifica: analisi bibliometrica comparata delle strutture di ricerca*. In: *La valutazione della ricerca e la valutazione della didattica*, a cura di Benito Vittorio Frosini. Milano: Vita e pensiero, 2011, p. 83-95, disponibile in *pre-print* a <<http://air.unimi.it/handle/2434/179445>>.
- [26] Sugimoto Cassidy R.; Pratt Jean A., *Using field co-citation analysis to assess reciprocal and shared impact of LIS/MIS fields*. «Proceedings of the American Society for information science and technology», 44 (2007), n. 1, p. 1-15, DOI: 10.1002/meet.1450440275.
- [27] Tarantino Ezio, *Troppo o troppo poco? Web of science, Scopus, Google scholar: tre database a confronto (un caso di studio)*, «Bollettino AIB», 46 (2006), n. 1-2, p. 23-33, <<http://bollettino.aib.it/article/view/5132>>, NBN: urn:nbn:it:aib-4881.
- [28] Turbanti Simona, *Cui prodest libraries authority work?*, «JLIS», 5 (2014), n. 2, p. 49-59, <<http://leo.cineca.it/index.php/jlis/article/view/10062>>, DOI: 10.4403/jlis.it-10062.
- [29] Wouters Paul F., *The citation culture*, tesi di dottorato, Università di Amsterdam, 1999, <<http://dare.uva.nl/document/487448>>.
- [30] Yang Kiduh; Meho Lokman I., *Citation analysis: a comparison of Google Scholar, Scopus, and Web of Science*, «Proceedings of the American Society for information science and technology», 43 (2006), n. 1, p. 1-15, DOI: 10.1002/meet.14504301185.

Articolo proposto il 16 agosto 2014 e accettato il 15 settembre 2014.

ABSTRACT AIB studi, vol. 54 n. 2/3 (maggio/dicembre 2014), p. 213-225. DOI 10.2426/aibstudi-10266.

SIMONA TURBANTI, Università di Pisa, Sistema bibliotecario di ateneo - Dottoranda presso il Dipartimento di Scienze documentarie, linguistico-filologiche e geografiche, Sapienza Università di Roma, viale Regina Elena 295, 00161 Roma, email simona.turbanti@uniroma1.it.

Navigare nel mare di Scopus, Web of science e Google Scholar: l'avvio di una ricerca sulla vitalità delle discipline archivistiche e biblioteconomiche italiane

Il contributo nasce parallelamente all'avvio di una ricerca all'interno del 29. ciclo di dottorato in Scienze documentarie, linguistiche e letterarie dell'Università La Sapienza di Roma, che si propone di analizzare il livello di salute degli studi del settore M-STO/08, inteso come capacità di uscire fuori dalla propria nicchia ed essere presenti in aree disciplinari e/o linguistiche esterne alla propria.

Si descrive la ricerca – effettuata nei due grandi database citazionali, Web of science e Scopus, e in Google Scholar – dei lavori scientifici dei docenti e ricercatori del settore M-STO/08, illustrando il metodo seguito, le principali differenze d'uso nonché i limiti dei database interrogati. Si analizzano in particolare i criteri adottati per l'organizzazione del lavoro in Scholar, il problema della corretta identificazione dell'autore e la successiva pulizia delle informazioni ottenute. Vengono quindi esposti i risultati raggiunti nell'interrogazione delle banche dati di Thomson Reuters e Elsevier e le difficoltà incontrate nell'ambiente Google.

Il quadro che emerge, oltre a confermare la complessità di uso degli strumenti citazionali, mostra come sia difficile ricavare risultati quantitativi rilevanti in contesti come quelli delle discipline LIS italiane, poco rappresentati nel panorama citazionale. La priorità emersa al termine di questa prima fase della ricerca è dunque quella della collocazione dei dati quantitativi all'interno di un corretta cornice di riferimento; solo in tal modo, infatti, si potrà cercare di rappresentare nel modo più oggettivo possibile i fenomeni di presenza e impatto – in altre parole, di internazionalizzazione e vitalità – delle discipline LIS italiane.

Navigating in Scopus, Web of science and Google Scholar: a research on Italian LIS studies vitality

This article has been conceived within the PhD in Library and Information Science at La Sapienza University of Rome, and is focused on the vitality and effectiveness of Italian LIS studies.

The paper describes the search in Web of science, Scopus and Google Scholar for any paper written by Italian LIS researchers and university professors, illustrating the research method, the difference of use – and limitations – of the three databases and the difficulties experienced with Google Scholar, and reporting WoS and Scopus numerical results as well.

This first research results made it clear that, in order to investigate the presence and impact of Italian LIS studies, a correct use and organization of quantitative data is essential.