

Analisi citazionale e indicatori bibliometrici di Antonella De Robbio

(per visualizzare il testo completo dell'articolo, cliccare qui)

1. La bibliometria

(cliccare qui per accedere alla versione integrale del testo)

La **bibliometria** è un ambito di studio applicato alle varie discipline scientifiche, che comprende anche l'**analisi bibliometrica**, sviluppatosi negli ultimi vent'anni grazie alla disponibilità online di banche dati di grandi dimensioni. La bibliometria, utilizza tecniche matematiche e statistiche per analizzare i modelli di distribuzione delle pubblicazioni e per esplorare l'impatto entro le comunità scientifiche. E' collocata entro un contesto più ampio di **scientometria**, la scienza per la misura e l'analisi della scienza.

2. Metodi di valutazione della ricerca

(cliccare qui per accedere alla versione integrale del testo)

La valutazione della ricerca richiede un duplice approccio:

- 1) **quantitativo**, ossia in termini numerici di impatto scientifico, brevettabilità, della presenza di contratti con aziende interessate a temi di ricerca;
- 2) **qualitativo**, ossia il giudizio dei pari (valutatori) o peer-review, che risulta a tutt'oggi il più importante dei metodi per una valutazione significativa della qualità. Appare evidente come sarebbe necessario associare ai metodi quantitativi uno o più metodi qualitativi.

Nell'analisi bibliometrica, le due misure bibliometriche più conosciute sono:

- il numero di citazioni;
- il fattore di impatto o impact factor IF, il numero di citazioni ricevute nell'anno corrente agli articoli pubblicati nei due anni precedenti diviso per il totale del numero di articoli pubblicati negli stessi due anni;

Oltre all'analisi bibliometrica (quantitativa), usata un po' ovunque, in altri paesi invece vengono usati più metodi, tra i quali il *panel* (modello anglosassone, ora dismesso dalla [RAE/2008](#)), il *peer review* (il giudizio dei pari) e l'analisi del beneficio economico.

3. Metodi bibliometrici

(cliccare qui per accedere alla versione integrale del testo)

L'analisi citazionale e *l'analisi di contenuto* sono metodi bibliometrici comunemente usati dalla bibliometria.

L'analisi citazionale, che usa citazioni nelle produzioni intellettuali scientifiche per stabilire connessioni ad altri lavori o altri ricercatori, è lo strumento cardine della disciplina di ricerca nota come bibliometria: essa è l'esame della frequenza e del modello di citazioni negli articoli e/o testi in generale. Sebbene per molti decenni il *Science Citation Index* dell'ISI [Institute for Scientific Information](#) - ora *Web of Science WoS* di Thompson è stato considerato lo strumento principe per la misurazione delle citazioni, da qualche tempo i servizi Web stanno mettendo in discussione il predominio non solo degli strumenti dell'ISI ma dello stesso IF.

L'analisi di contenuto invece, nota anche come *analisi testuale* quando condotta esclusivamente su testi, è una metodologia standard nel campo delle scienze sociali applicata per lo studio del contenuto della comunicazione. In tale contesto per analisi dei dati o *data analysis* si intende l'atto di trasformare un dato con lo scopo di estrarne informazioni utili e facilitare conclusioni che abbiano un significato logico.

4. Indicatori bibliometrici

(cliccare qui per accedere alla versione integrale del testo)

L'**Impact Factor** (IF) è l'indicatore bibliometrico più conosciuto nell'analisi bibliometrica. Tuttavia, esso non è il solo indicatore proposto dall'ISI (ora Thompson) a cui fare riferimento. Altri indicatori utilizzati sono:

- 1) **Immediacy Index**: misura quanto successo sta avendo il lavoro nell'anno di pubblicazione e in relazione a quanto velocemente un articolo della rivista è mediamente citato e quanto spesso gli articoli della rivista sono citati nello stesso anno;
- 2) **Cited Half Life**: misura la validità nel tempo degli articoli citati o la durata delle citazioni nel tempo;
- 3) **Rate of Cites Index**: rappresenta un indice di qualità del singolo lavoro, basato sull'assioma che quanto più il lavoro è citato da altri ricercatori tanto più rilevante è il suo valore scientifico;
- 4) **Citation Impact**: è calcolato per uno specifico soggetto o autore o istituzione o paese sulla base del rapporto tra il numero di citazioni ricevute e il numero di articoli pubblicati.

5. L'impact Factor (IF)

(cliccare qui per accedere alla versione integrale del testo)

L'**Impact Factor** (IF): misura il livello della ricerca scientifica, su scala nazionale ed internazionale, delle pubblicazioni scientifiche. Varia moltissimo da area ad area; è, infatti, una misura del numero di citazioni dei lavori pubblicati in una certa rivista rispetto al numero totale di lavori pubblicati dalla stessa rivista negli anni precedenti.

6. Principali difetti dell'Impact Factor

(cliccare qui per accedere alla versione integrale del testo)

L'IF è un parametro di riferimento che consente di stilare graduatorie delle riviste per ciascuna delle categorie disciplinari adottate dal *JCR*, [Journal Citation Report](#). Anche se generalmente accettata, l'assunzione che ad un'elevata frequenza di citazione corrisponda un'elevata qualità della rivista, sul piano concettuale, ha scarso fondamento. Questa pratica, grazie alla quale l'impact factor viene usato come parametro di riferimento per stilare graduatorie delle riviste più prestigiose, è stata oggetto di una serie di critiche avanzate in particolare da alcuni matematici, in primis Alessandro Figà Talamanca nel suo storico discorso dal titolo [L'Impact Factor nella valutazione della ricerca e nello sviluppo dell'editoria scientifica](#) al convegno SINM del 2000.

7. Indici bibliometrici alternativi all'Impact Factor

(cliccare qui per accedere alla versione integrale del testo)

Il Web ha fatto sorgere centinaia di nuovi database o strumenti che permettono la ricerca di citazioni. Questi strumenti ricadono entro tre categorie:

- Strumenti o database che permettono all'utente di cercare nel campo a testo pieno per determinare se un certo item, autore o periodico, sono stati citati in un documento: [arXiv.org](#) (eprints server), [Google Book Search](#) (progetto digitalizzazione testi), [Journal of Archive](#) di IP, [Physical Review Online Archive](#), [Scirus](#) di Elsevier (motore di ricerca), [CiteSeer](#) (automated o autonomus citation indexing ACI), [Google Scholar](#) (motore di ricerca con sistema di ranking basato su ACI).
- Strumenti o database che permettono all'utente di ricercare nel campo dei riferimenti citati (articoli che citano altri articoli) per identificare citazioni rilevanti: [MathSciNet](#), [Science Direct](#) di Elsevier, [SciFinder](#), [Scholar](#) di CAS, [SPIRES-HEP](#) (eprints server di Stanford), [Scitation/SPIN](#) di AIP, [NASA Astrophysics Data System](#).
- Database che lavorano esattamente come Web of Science: il principale esempio di questa categoria è [Scopus](#). Recentemente, Scopus ha rilasciato un nuovo servizio gratuito [TopCited](#) che permette di consultare gli articoli più citati nelle varie discipline usando un'API integrata con Google Map. E' interessante notare come i 20 articoli più citati negli ultimi 3-5 anni in tutte le discipline (26 aree di soggetto) sono ad accesso aperto.

8. L'Eigenfactor (EI)

(cliccare qui per accedere alla versione integrale del testo)

L'[Eigenfactor](#) è un nuovo modello matematico per calcolare l'impatto delle riviste scientifiche accademiche che tiene conto dei diversi standard di citazione tra una disciplina e l'altra. Esso rappresenta un nuovo metodo per la valutazione dell'influenza (impatto) dei periodici accademico-scientifici ai fini di ottenere una mappatura della struttura della ricerca accademica.

9. L'Indice di Hirsch o H-index e sue varianti

(cliccare qui per accedere alla versione integrale del testo)

Con **h-index** si intende un indice proposto nel 2005 da Jorge E. Hirsch della University of California di San Diego per quantificare la prolificità e l'impatto del lavoro degli scienziati, basandosi sul numero delle loro pubblicazioni ed il numero di citazioni ricevute. Esso assume grande rilevanza poiché verifica la reale influenza di uno scienziato sulla comunità, prescindendo da singoli articoli di grande successo, o anche dai lavori di autori che pur avendo pubblicato molto, hanno prodotto solo articoli di scarso interesse, come invece avviene usando l'IF. I calcolatori di H-index sono facilmente reperibili in rete e sono accessibili a chiunque: [scHolar INDEX](#), per esempio, è un software che interroga Google Scholar tramite queries e fornisce H-index dell'autore e/o della pubblicazione richiesta per verificare l'impatto dell'opera nella comunità scientifica in generale. Un altro software che consente il calcolo dell' H-index su uno specifico autore è [QuadSearch – MetaSearch Engine](#).

Altra variante dell'indice H è l' **h-b-index (Hirsch-Banks Index)**, meglio un'estensione dell'indice H sviluppato da Michael Banks del Max Planck Institute for Solid State Research, focalizzato sulle pubblicazioni in fisica dello stato solido. E' definito allo stesso modo dell'Indice H, ma è basato sulla ricerca di uno specifico argomento o composto, invece che del nome dello scienziato.

Il curatore di *Science Focus*, Jin Bihui ha inventato un ulteriore metodo, detto "**a-index**" il quale è definito come la media del numero di citazioni ricevute da opere nel numero di pubblicazioni "h-index".

[Publish or Perish](#) (PoP) è un software scaricabile gratuitamente dal sito sul proprio pc che recupera e analizza le citazioni accademiche e consente di calcolare una serie di indici bibliometrici di uno specifico autore: oltre ai classici *h-index* e *g-index*, esso permette di calcolare anche il *contemporary h-index* (hc-index), l'*individual h-index* (hI-index) e il *Age-weighted Citation Rate* (AWCR). Publish or Perish, inoltre, ha implementato un metodo alternativo individuale sul modello H-index, chiamato *Normalized Individual h-index* (hI,norm), con un differente approccio: invece di dividere l'h-index totale, prima normalizza il numero di citazioni per ogni paper dividendo il numero di citazioni per il numero di autori che contribuiscono al paper e poi calcola l'h-index sul conteggio normalizzato delle citazioni.

[SCImago Journal & Country Rank](#) (SJR) è un nuovo database, lanciato nel dicembre dello scorso anno, che è accessibile in Internet gratis e che permette di generare statistiche sulle citazioni degli articoli pubblicati nelle riviste peer-reviewed. Questo strumento genera delle statistiche per paese e confronta il numero degli articoli pubblicati, le citazioni, per paese.

10. La metrica del Web e il Web Impact Factor (WIF)

(cliccare qui per accedere alla versione integrale del testo)

Almind e Ingwersen, considerati tra i più accreditati studiosi di metodi bibliometrici, hanno pubblicato nel 1997 i primi studi relativi all'area di problemi definita con il termine **Webometrics**. La scienza detta Webometrics o Webmetrics o anche cybermetrics, o web metrics (metrica del web) tenta di misurare il World Wide Web al fine di ottenere conoscenze sul numero e sul tipo di connessioni ipertestuali gli hyperlinks, sulle strutture del Web e sui modelli di utilizzo delle risorse che popolano il cyberspazio.

Si tratta dello studio degli aspetti quantitativi della costruzione e dell'uso delle risorse informative delle strutture e delle tecnologie sul web valutate secondo un' approccio bibliometrico ed informetrico (Björneborn & Ingwersen, 2004). il **Web Impact Factor** WIF è basato sull'analisi dei link e si ricava dal numero dei contatti che un documento pubblicato sul web riceve moltiplicato per il numero delle pagine.

11. Iniziative per la standardizzazione delle statistiche e per la misurazione dei dati di utilizzo

(cliccare qui per accedere alla versione integrale del testo)

Se gli indicatori generati dall'autore sono le citazioni, quelli generati dai lettori sono i dati sull'utilizzo. Questi dati possono essere raccolti tramite server web e *linkresolver logs*. E' necessario però normalizzare il lavoro di raccolta dei dati di log al fine della condivisione di metodi per ottenere analisi significative.

Al recente [workshop OAI5](#) dello scorso aprile 2007 alcuni interventi si sono focalizzati proprio sulla misurazione delle risorse entro gli archivi aperti e risorse Open Access. Johan Bollen del Digital Library Research & Prototyping Team presso la biblioteca del Los Alamos National Laboratory Research ha presentato il programma *MESUR metrics from scholarly usage of resources*¹, che offre un set di strumenti per la valutazione degli oggetti della comunicazione scientifica.

Esistono iniziative volte a fornire infrastrutture adeguate che includono i dati dell'editore e i dati raccolti attraverso gli archivi ad accesso aperto. La più famosa tra queste è [COUNTER](#) (*Counting Online Usage of Networked Electronic Resources*),

12. L'Usage Factor (UF)

(cliccare qui per accedere alla versione integrale del testo)

L'*Usage Factor UF*, definibile come un indicatore bibliometrico quantitativo di nuova generazione, è complementare al tradizionale modello noto dell'*Impact Factor IF*.

L'UK Serials Group (UKSG) ha commissionato ricerche entro un progetto che ha lo scopo di sondare la possibilità di utilizzo di un Usage Factor come indice per misurare il valore e l'autorevolezza delle riviste scientifiche e delle pubblicazioni accademiche in generale.

In sostanza il progetto mira a sondare la fattibilità reale di uno sviluppo di un nuovo indice bibliometrico di misurazione del valore dei periodici scientifici basato sull'utilizzo delle risorse elettroniche che potrebbe essere appunto definito come Usage Factor o Fattore di Utilizzo.

¹ Relazione adOAI5

[http://indico.cern.ch/getFile.py/access?
contribId=6&sessionId=14&resId=1&materialId=slides&confId=5710](http://indico.cern.ch/getFile.py/access?contribId=6&sessionId=14&resId=1&materialId=slides&confId=5710)