



EDITORIALE

DORA: una critica radicale all'Impact Factor. Come valutare e misurare la qualità dei risultati della ricerca con metodi alternativi?

The case of DORA: a radical criticism of the Impact Factor. How we may assess and measure the research output quality, with alternative methods?

Ciappelloni Raoul

Abstract. In this editorial the San Francisco "Declaration on Research Assessment" (DORA, 2012) is discussed. The initiative widely criticizes the use (and abuse) of metrics, such as the Journal Impact Factor (IF), now widely used in order to assess the "scientific value of scholarly Journals" and the researchers themselves. In particular, the IF, over the time, has become a sort of convenient scholar benchmark. For this reason the advantages and disadvantages of the "IF" are outlined in the paper. Actually, a part of the international scientific community, considers IF and similar methods, an "unacceptable simplification". DORA, comes in the midst of this heated international debate; it propose the search for alternatives methods that can take into the account the real value of a scientific research output. This hope, however, still needs to find effective implementation. It seems to be still a long way to go, to find a solution

Riassunto. In questo editoriale, parleremo della "Declaration on Research Assessment" (DORA), di San Francisco (2012) una iniziativa che critica il Journal Impact Factor (IF), largamente utilizzato per stimare il valore delle riviste scientifiche e gli stessi ricercatori. La critica espressa con questa iniziativa è che le metriche basate su indici, in particolare l'IF, nel tempo, sono diventate una specie di benchmark usato per misurare le performance della ricerca. Per questo motivo nell'editoriale sono discussi vantaggi e svantaggi dell'IF. In effetti, una parte della comunità scientifica internazionale considera questo e simili metodi, una semplificazione inaccettabile. DORA, arriva nel pieno di questo acceso dibattito internazionale e intende proporre alternative che possano tenere conto dell'effettivo valore scientifico di un output della ricerca. Tale auspicio tuttavia deve ancora trovare effettiva realizzazione. Sembra che a tal riguardo la strada da percorrere non sia breve

DORA: L'antefatto

In questo editoriale, parleremo diffusamente di DORA, ovvero della "Declaration on Research Assessment - Dichiarazione sulla valutazione della ricerca", una iniziativa che coinvolge l'IF (o Journal IF; sta per Journal Impact Factor), l'indice più discusso e tuttavia largamente utilizzato per stimare il valore o, in modo più evocativo, "l'impatto" delle riviste scientifiche sulla comunità degli studiosi (e, a seguire, delle istituzioni di ricerca e degli stessi ricercatori).

La cosa sembra non piacere a un numero crescente di scienziati, che prende apertamente una posizione critica, anche perché l'IF, nel tempo, è diventato una specie di benchmark, pervasivo e forz'anche sbrigativo, usato per misurare un po' di tutto: dall'importanza di una pubblicazione, al valore di un progetto di ricerca.

Nell'editoriale affronteremo in particolare le motivazioni di quella parte della comunità scientifica internazionale che considera l'uso di metriche di impatto, una semplificazione inaccettabile.

Partiamo dalla scontata considerazione che, in senso generale, il valore scientifico (al pari della bellezza di una persona o della bontà di un cibo), è un dato che, per sua natura, esprime una qualità.

I sostenitori dell'IF avrebbero invece la pretesa di manifestarlo (il valore scientifico)

attraverso una pura quantità, addirittura un numero, in modo che possa essere comparato con altri valori numerici o inserito in formule. È per tale motivo di fondo che l'approccio mostra problemi di interpretazione e la sua semplice applicazione non è accettata (ne forse accettabile) in ogni disciplina.

Il "padre" dell'IF, come oggi lo conosciamo, è Eugene Garfield. Si tratta di uno studioso, di Chimica, linguistica, biblioteconomia. Proprio in quest'ultima veste, lo aveva detto subito che il metodo era stato ideato per valutare l'opportunità di acquistare una rivista scientifica rispetto ad un'altra e facilitare la vita dei bibliotecari.

Principalmente per questo motivo, l'indice ha goduto di ampia popolarità fra gli Editori e negli Enti di ricerca, riuscendo a ottimizzare l'impiego dei fondi destinati agli abbonamenti. La cosa è stata provvidenziale in un trend di finanziamenti pubblici decrescenti e costi delle fonti documentali lievitanti (un fenomeno causato dallo strapotere monopolistico delle case editrici scientifiche).

Ad un certo punto però e non senza contrasti, l'IF cominciò ad essere utilizzato anche per definire il valore dell'attività del singolo ricercatore, in virtù di una specie di sillogismo approssimativo: "Se A è una rivista con alto impatto e B pubblica un articolo su A; allora l'articolo di B ha un alto impatto".

Ma la conclusione del ragionamento è, con ogni evidenza, debole. Potrebbe infatti accadere che, per varie ragioni, un articolo fiacco e scarsamente significativo, venga pubblicato, su una rivista prestigiosa (cioè citata). L'articolo in questione, pur trovandosi collocato nel "posto giusto", non farà mai "tendenza" e la comunità scientifica lo ignorerà. Nonostante ciò, al suo autore, almeno formalmente, verrà attribuito un certo immeritato prestigio, in virtù della pura "localizzazione" del lavoro e non del suo effettivo "valore" (scientifico). Non si tratta di una storia nuova. Tredici anni fa, il Professor Alessandro Figà Talamanca, in una comunicazione presentata al IV Seminario sul sistema informativo nazionale per la Matematica (Lecce, 2 Ottobre, 2000) e ormai diventata un "classico"; ebbe modo di dire tutto il male possibile dell'Impact Factor. I suoi dubbi, da allora, vengono di tanto in tanto "riscoperti" dai ricercatori o da semplici commentatori, che criticano con veemenza, qualche aspetto dell'IF. Sarà forse perché si parla di valutazioni che sono sempre un po' antipatiche? Il sospetto c'è. A parte questo, non si tacere l'esistenza di alcuni punti effettivamente oscuri dell'Indice in questione, che riassumiamo brevemente di seguito:

- L'IF: è: *scarsamente indicativo del vero valore scientifico*. Si può dire che sostituisca una valutazione qualitativa "a tutto tondo", fatta da "umani" con un secco indice quantitativo elaborato in modo semi-automatico. Di certo quest'ultimo non potrà tener conto delle particolarità (positive e negative) dell'attività scientifica, per non dire letteraria, di uno studioso;

- L'IF è: *buono per alcune discipline e non per altre*. C'è una innegabile influenza dell'ambito disciplinare e delle sue caratteristiche intrinseche, sulla affidabilità della valutazione fatta con indici scientometrici. Non dimentichiamo che l'IF nasce e prospera nel mondo biomedico. Invece nell'ambito umanistico dove, per dirne una, le monografie sono ancora più rappresentative dei singoli articoli, questo approccio si dimostra applicabile solo con determinati aggiustamenti (cosa che induce alcuni autorevoli studiosi, a proporre di non applicarlo affatto);

- L'IF è: un "*depressivo*" editoriale. I ricercatori che possono collocare i propri articoli su riviste con eccellente IF, in genere, rifuggono i Journal con basso (o nullo) impatto, trovandoli poco attrattivi (normalmente si tratta di riviste non scritte in lingua inglese, non localizzate negli USA e non prioritariamente collegate al Web).

Dovendo pur scriverci qualcosa, è lì che si "sparano le cartucce di minor valore", perché essendo la rivista poco citata, non porterà loro grande prestigio e vantaggi (leggi grande IF). Così, nel tempo, questa situazione ha contribuito a diminuire la biodiversità editoriale, penalizzando Journal scientifici, anche validi, ma non recensite da JCR (Thomson) o da Scopus (Elsevier).

È probabilmente per questo (oltre naturalmente che per l'alta qualità della ricerca), che gli Stati Uniti detengono il massimo numero di riviste con alto IF (2806 riviste. Il più alto IF è di 153.459). Tra parentesi, osserviamo che l'Italia (il numero di riviste è 127. Il maggior IF è di 8.974, come riportato da Thomson Reuters©; 2012, JCR Science Edition), rappresenta un "mondo di mezzo", non però così disonorevole se paragonato ad altri Paesi europei.

- L'IF è: *potenzialmente truffaldino e limitato*. Il calcolo di detti indici non sarebbe proprio trasparente e, pertanto, il metodo si presterebbe (almeno teoricamente) a piccoli ma significativi aggiustamenti da parte degli stessi autori, definiti da Figà Talamanca: "allegra brigata di entusiasti reciproci-citanti" (Talamanca, 2000). Ma non basta. Oltre a ciò, dovremmo anche considerare che esiste una vasta "letteratura" scientifica, di grande valore, che però non genera alcun impatto, non rientrando in questi meccanismi di calcolo. Pensiamo ai dataset, al software, all'attività didattica, per non citare quella legata ai brevetti. Sarebbe però riduttivo considerare tutto ciò come una faccenda in fondo interna all'esoterico mondo della scientometria.

L'affaire DORA, arriva nel pieno di un dibattito internazionale sul finanziamento della ricerca e va al centro di una questione, quella della meritocrazia nella scienza, trattata da intere istituzioni (ricordiamo il RAE nel Regno Unito - <http://www.rae.ac.uk/> e l'ANVUR in Italia - <http://www.anvur.org/>) e del loro piccolo esercito di professionisti valutatori. Si tratta poi anche di un confronto fra interessi contrapposti di soggetti diversi:

- Cultori di discipline biomediche e Studiosi afferenti al settore umanistico,
- Ricercatori ed Editori,
- Studiosi istituzionalizzati e Studiosi precari. In ultimo, potremmo semplicemente dire: valutatori delle performance e coloro che di queste valutazioni pagheranno le conseguenze.

L'IF è sembrato e sembra tutt'ora a molti, uno strumento con il quale si vorrebbe risolvere in modo un po' sbrigativo, l'intricato problema della valutazione dell'attività di ricerca. Un po' come fece Alessandro Magno a Gordio, dove invece di provare a dipanare il famoso nodo, con un colpo di spada lo tagliò mostrando, se non pazienza e abilità, almeno un atteggiamento assai pratico.



Alessandro il Grande taglia il nodo Gordiano,
di Giovanni Paolo Panini, (olio su tela, 1718 - 1719)
[immagine di pubblico dominio]

Vi sono quindi diverse sensibilità in campo. Forse anche per questo, la discussione non ha sempre mantenuto toni pacati. È addirittura finita sotto i riflettori della stampa generalista italiana, così poco incline ad affrontare le questioni scientifiche (vedi: "Gli insorti della DORA" di Sylvie Coyaud, Blog della Repubblica, 17 maggio 2013 - <http://ocasapiens-dweb.blogautore.repubblica.it/2013/05/17/gli-insorti-della-dora/>; come anche: "Impact Factor con giudizio" di Vittorio Lingiardi su Il Sole 24 Ore 30 Giugno 2013, http://www.scienzainrete.it/files/Sole_2013_06_30_Lingiardi.pdf).

Ha preso così vita una discussione che in Italia, per contiguità e analogie come con il famoso test delle macchie (di Rorschach), ha inevitabilmente finito per evocare ben "altri" temi presenti nel confronto sociale e politico: efficienza della spesa pubblica, legittimazione del sistema Statale (universitario e non solo). Forse perché, anche collettivamente, è ormai questo che abbiamo in mente e che spunta fuori in ogni occasione.

DORA in due parole

È giunto il momento di dire qualcosa di più sulla declaration. Per questo interpreteremo alcune interessanti considerazioni contenute in un editoriale di Stefano Bertuzzi e David Drubin, apparso recentemente su *Molecular Biology of the Cell* (15 Maggio 2013; <http://www.molbiolcell.org/>). I due scienziati sembrano voler giocare con i sottintesi. Infatti, per spiegare la loro tesi "anti IF" immaginano di dar vita a un ipotetico progetto valutativo, chiamato Progetto MESS (l'acronimo sta per: Metric Evaluation for Scientific Scholarship; ma la parola "mess" in inglese significa: casino, confusione, bordello). Come suggerisce il facile doppio senso, si tratta di un sistema-bufala, ma viene discusso seriamente dai due studiosi, per mettere in evidenza le carenze degli approcci basati sugli algoritmi che portano a indici.

Mess viene presentato come un productivity index, un sistema automatico (o almeno altamente automatizzabile) basato su un algoritmo in grado di stimare l'attività di ricerca tenendo conto di diversi fattori come: il numero e l'importo dei finanziamenti attribuiti a ogni ricercatore, il numero di pubblicazioni da lui firmate, i download e le citazioni che le hanno riguardate. Oltre a ciò, si prevedono appositi aggiustamenti numerici per tener conto anche

del prestigio dei giornali in cui gli articoli sono stati pubblicati.

In tal modo finalmente sarebbe possibile: "... stilare una classifica di tutti i ricercatori nel mondo, fornendo una base per l'assunzione, il finanziamento e decisioni relative alla carriera".

Detto così, sembrerebbe un'ottima cosa.

Ma i due ricercatori si affrettano ad aggiungere: " ... *Tale spaventoso sistema, avrebbe alcuni effetti collaterali*". Quali sarebbero? Nell'editoriale si dice lapidariamente che in tal modo si favorirebbero i ricercatori più anziani che lavorano nei campi di studio maggiormente finanziati e ricchi di pubblicazioni, sostituendo "... *la ricerca di una maggiore comprensione (scientifica) con la ricerca di sempre più MESS*".

Quindi Bertuzzi e Drubin suggeriscono che i sistemi "IF-like" finiscano con il privilegiare una parte della comunità scientifica (e una particolare modalità, o stile, di lavoro), a scapito dell'altra.

Oltre a ciò, verrebbe avviato un perverso feedback, per cui il sistema MESS si alimenterebbe ricorsivamente dei suoi stessi dati, peggiorando via via la situazione, così che tutto alla fine precipiterebbe nella confusione, appunto, in un grande "mess". La conclusione del ragionamento è:

"... Per quanto assurdo possa sembrare questo immaginario modo di valutazione, tristemente, la comunità di studiosi fa affidamento su un sistema che è almeno tanto assurdo quanto quello ipotizzato e i ricercatori possono prendersela solo con se stessi per questo sfavorevole stato delle cose".

Cominciamo a dire che gli editorialisti sono autorevoli. Stefano Bertuzzi è Direttore dell'Office of Science Policy, Planning, and Communications al National Institute of Mental Health (NIMH, uno dei National Institutes of Health degli Stati Uniti), recentemente è stato nominato dall'American Society for Cell Biology: Executive Director, mentre David Drubin è Professore di "Cell and Developmental Biology" all'Università di Berkeley ed editor in chief della rivista *Molecular Biology of the Cell*.

Dobbiamo creder loro? Le argomentazioni sono convincenti e la tesi che essi sostengono riporta una forte amplificazione, con alcune varianti, dei difetti conosciuti dell'IF. Infine il loro MESS è una parodia realistica di quanto alcuni autorevoli ricercatori pensano si dovrebbe fare per realizzare un buon sistema, basato su logica quantitativa, di valutazione scientifica.



San Francisco Declaration on Research Assessment (DORA), iniziativa avviata dall' American Society for Cell Biology (ASCB) con un gruppo di autori ed editori di giornali scientifici

L'eccellente editoriale da maggiore valore al contenuto della Declaration di San Francisco,

tuttavia non risolve la questione sul tappeto, come potremo constatare con un'occhiata più da vicino al documento.

Diciotto raccomandazioni

Il nocciolo di DORA è contenuto in un elenco di 18 raccomandazioni (<http://am.ascb.org/dora/>), promulgato, nel corso dell'Annual Meeting di San Francisco, dell'American Society for Cell Biology - ASCB (<http://www.ascb.org/>; 16 Dicembre 2012). Un po' come nel caso delle Novantacinque Tesi di Lutero si argomenta, con grande passione e pubblicamente, contro quello che viene visto come un potente establishment (anziché teologico, questa volta è scientifico-editoriale) e lo si fa, per far cessare una deprecabile modalità, da questi introdotta a proprio esclusivo vantaggio.



Martin Lutero affigge le sue Novantacinque Tesi sulla porta della chiesa di Wittenberg nel 1517
[dipinto di Ferdinand Pauwels (1872). Immagine di pubblico dominio]

Però, prima di procedere, è giusto rendere omaggio alla coerenza, riportando un fatto. La rivista "Molecular Biology of the Cell", organo di stampa dell'ASCB, che ha iniziato tutta questa querelle, al contrario di ciò che fanno gli altri journal scientifici, omette zelantemente di menzionare il suo Fattore di Impatto (che non è oltretutto basso. Il JCR®, per il 2012, riporta un valore di 4.604), per cui all'interno della sua pagina di Welcome, come pure nel resto del materiale pubblicato, non ne troverete traccia.

La Declaration è stata firmata in prima battuta da una ottantina di istituzioni. Fra loro figurano nomi di rilievo come: Public Library of Science (PLOS), The EMBO Journal, FEBS Letters ed altre.

Successivamente si sono aggiunte oltre trecentosettanta Istituzioni (situazione aggiornata alla fine di Luglio 2013), fra cui alcune del nostro Paese (Fondazione Telethon, Società italiana di filosofia morale, Italian Association of Psychology (AIP), SIBBM, Italian Society for Biophysics and Molecular Biology, Veneto Institute of Oncology, IOV-IRCCS, Associazione Italiana di Cultura Classica, The Italian Society for Horticultural Science, STOREP - Associazione Italiana per la Storia dell'Economia Politica).

Anche a livello personale, le adesioni da tutto il mondo sono state numerosissime (più di novemila) mostrando che questa problematica è davvero molto sentita.

L'analisi sintetica della tipologia di istituzioni aderenti si può fare partendo dai dati presenti nello stesso sito dell'ASCB. Bastano due conti per dire che il 94% degli Enti partecipanti sono

affendenti alle discipline scientifiche e il 6% a quelle umanistiche. Inoltre la maggior parte delle istituzioni proviene da Europa e Stati Uniti (83.6%), piuttosto distanziato il resto del mondo (Asia, Medio Oriente, Africa e altre, 16.3%), un quadro che di per se è significativo e mostra come questa problematica sia probabilmente più sentita nelle zone geografiche dov'è maggiormente attiva l'editoria scientifica.

Il documento con le 18 raccomandazioni è articolato in sei titoli che denotano altrettanti brevi paragrafi. Il primo è una General Recommendation che esprime il leitmotiv di tutto il documento e consiglia di: *"Non usare metriche journal-based, quale l'IF, come rappresentative della qualità dei singoli articoli di ricerca, per stabilire il valore del contributo di ogni singolo scienziato, o per deciderne l'assunzione, promozione o finanziamento"*.

Seguono raccomandazioni più specifiche per:

- Agenzie di finanziamento della ricerca,
- Istituzioni,
- Editori,
- Organizzazioni che forniscono modalità di valutazione,
- Ricercatori.

Più che indicazioni su come effettuare una valutazione della ricerca con modalità alternative, qui si parla di atteggiamenti auspicabili in senso più generale. Si dice, ad esempio, di esplicitare le metodologie di valutazione dei titoli scientifici di merito, cercare di tenere conto di tutta la produzione scientifica di un autore e non soltanto di un subset rappresentativo. Al fine di facilitare il processo di valutazione dei candidati, si consiglia inoltre di specificare meglio possibile le caratteristiche di merito per ottenere un finanziamento o una promozione. Si arriva a suggerire (nelle raccomandazioni rivolte agli editori), di ridurre il più possibile l'enfasi sul fattore di Impatto come mezzo di promozione della rivista scientifica o, almeno, di presentare contestualmente un certo numero di metriche alternative (5-year Impact Factor, EigenFactor, SCImago, h-index) e così via.

Sono citate anche questioni che, pure auspicabili, non sembrano direttamente collegate alla stima del valore scientifico. Ad esempio si consiglia un uso maggiore delle licenze Creative Commons (<http://creativecommons.org/>). Viene anche suggerito agli autori, di preferire la citazione degli articoli originali rispetto alle reviews, che generano una artificiosa amplificazione dei riferimenti bibliografici.

La "raccomandazione numero 17", è rivolta ai ricercatori. In essa si dice: "Quando impegnati in comitati di valutazione accademica che prendono decisioni circa finanziamento, assunzione, attribuzione di posizioni accademiche, si faccia in modo che i giudizi siano basati sul contenuto scientifico delle pubblicazioni, piuttosto che sulle loro metriche di impatto". È l'espressione di una frontale sfiducia nelle attuali modalità di valutazione degli articoli scientifici, non importa se passati al vaglio di una revisione paritetica. La conclusione cui si giunge al termine di ogni sezione del documento è: "fare ogni sforzo per evitare di basarsi sul

Journal Impact factor, ma promuovere metodologie che pongano in luce l'autentico valore e influenza dei portati della ricerca".

Ma quanti la pensano così? La letteratura citata al termine del documento è scarna e un po' datata (i lavori vanno dal 1997 in poi, e solo uno è del 2012). Ciò indica che la discussione sull'argomento, pur essendo significativa, non ha ancora portato a una "massa critica" di contributi e proposte. Siamo quindi in una fase iniziale in cui però, i contenuti della Dichiarazione, sono già strutturati e, in linea di principio, condivisibili. Essi esprimono l'esigenza, generalmente avvertita da "scientists" appartenenti ai più diversi ambiti disciplinari, di migliorare le modalità di valutazione non solo delle pubblicazioni ma, come ben si dice, degli "output" della ricerca, potendo questi essere di varia tipologia. Certo in questo momento, nel suo insieme, la genericità della Declaration di San Francisco è disarmante. Non c'è neppure la vaga indicazione operativa sulle alternative all'IF. Il fattore di impatto non piace e basta. Ma cosa significherebbe farne a meno? Molti si provano a dare delle ricette. Un metodo viene delineato da una Company chiamata ImpactStory(<http://impactstory.org/>), che troviamo nel gruppo dei primi firmatari di DORA. Si tratta di una iniziativa piuttosto "giovane" se pensiamo che è nata nel 2011 in seno al "Beyond Impact Workshop" tenuto a Londra (<http://beyond-impact.org/>) ed è sostenuta dalla Fondazione Alfred P. Sloan di New York, quindi da fondi privati.

ImpactStory è dedicata alla valutazione dell'impatto della ricerca con metriche alternative all'IF (definite "Altmetrics". Vedi l'hash tagt #altmetrics in Twitter®) ispirandosi fondamentalmente alla filosofia Open Source. Il metodo è Human-centered e di natura "composita", capace di tenere conto del valore delle varie componenti dell'attività scientifica, come: pubblicazione di articoli, di libri, data set, presentazioni di slide, produzione di dati e software. Tutto fa brodo e ogni output della ricerca viene compreso nell'analisi globale. Ovviamente il metodo è dipendente dalla disponibilità di tools e servizi Web (Orchid, Google Scholar, GitHub, Slideshare, Twitter, Delicious).

Ad esempio, un articolo potrebbe essere valutato in termini di quanto è: citato da Google Scholar, discusso nei social network, indicato e messo a disposizione per il download dagli utenti del Web e via dicendo. Così, chi sia effettivamente interessato, e abbia tempo a disposizione (questo sembra essere il problema), può arrivare a farsi un'idea dell'attività scientifica in questione in termini di innovatività, avanzamento, ed altro. (consigliamo di visionare l'interessante materiale online e soprattutto le FAQ: <http://impactstory.org/faq>).

A tal riguardo, due considerazioni:

(1) Si può osservare che il metodo Impactstory non sembra fornire la possibilità di stilare una ordinaria semplice graduatoria, poiché inevitabilmente il giudizio ritorna di natura qualitativa. Ad esempio, stabilendo che alcuni punti di un certo profilo, siano prioritari, gli altri saranno considerati meno significativi, opinione che potrebbe essere però ribaltata da un altro valutatore. Sarebbe allora giocoforza valutare (o pesare) lo stesso giudice, cosa che porterebbe a una situazione un po' surreale. (2) Inoltre, per la forte dipendenza dalla Rete, ben difficilmente si potrebbe valutare in questo modo l'attività di uno studioso di area umanistica, dove questi tools Web based sono un po' meno significativi.

In futuro, il rating dell'attività scientifica potrebbe anche diventare un'area di business di qualche interesse. Lo dimostra l'agenzia informativa Thomson Reuters, che pubblica una versione dell'IF, altri indici derivati, grafici esemplificativi e dati nell'ISI Web of Knowledge (<http://thomsonreuters.com/web-of-knowledge/>), fornendo strumenti di analisi in modo non molto diverso da quello che fanno le società di ratings, come Standard and Poor's (<http://www.standardandpoors.com/>). Thomson non si spinge certo a decretare un "default", o

inaffidabilità scientifica di un soggetto, ma fornisce la possibilità di farlo. Tra l'altro, Standard and Poor's è, guarda caso, parte del gruppo McGraw Hill, editore scientifico di primo livello che ha lanciato nel 2009 in collaborazione con JAMA (una casa editrice di giornali medici), l'iniziativa "Online Evidence-Based Medicine, Clinical, and Educational" per studenti, docenti e clinici.

C'è poi anche un vero editore di riviste scientifiche, come Elsevier, che non si è limitato a diffondere pubblicazioni, ma si è messo a fornire proprie valutazioni scientometriche con Scopus® (<http://www.elsevier.com/online-tools/scopus>) entrando così, in prima persona, nell'intricata vicenda.

In un recente documento del Giugno del 2013 (<http://elsevierconnect.com/san-francisco-declaration-on-research-assessment-dora-elseviers-view/>), l'editore specifica che: "... *pur non sottoscrivendo DORA nella sua interezza...*", esprime sostegno all'iniziativa, anche se unicamente in relazione a: "...*quegli elementi di DORA che riflettono i problemi ben noti dei fattori di impatto, in relazione ai quali abbiamo sostenuto una ampia gamma di alternative e di best practices*".

Notiamo che se uno dei massimi esponenti dell'editoria scientifica (legato alla logica dell'IF), dice questo, ciò significa che nel contesto del Research Assessment bisogna proprio "esserci", starci dentro, giocare un ruolo, in questo caso di partecipazione e in fondo di controllo, sia pure effettuato in modo corretto e per fini di informazione e orientamento. Perché qui si gioca davvero il futuro del sistema basato sul finanziamento pubblico dell'attività di ricerca.

Potremmo azzardare in chiusura una specie di provocazione propositiva. Se Dora esprime l'idea che i metodi scientometrici siano scarsamente utili per valutare l'eccellenza nella ricerca, sarebbe allora forse possibile impiegarli per valutare la sua mediocrità. Si potrebbero impostare dei "limiti" per ogni disciplina, cioè dei bassi valori di IF, sotto i quali un candidato non viene preso in considerazione per la posizione o per l'attribuzione del finanziamento. Eliminando una parte dei concorrenti sarebbe più agevole l'applicazione dei metodi di giudizio "qualitativi" e time consuming sui restanti. Questi potrebbero fruire dell'analisi lunga e dettagliata di cui si parlava prima; analisi effettuata da revisori che alla fine condurrebbero a una fase "documentabile" e condivisa.

Quest'ultimo è un punto davvero importante e forse poco considerato. Se infatti è necessario che "... *i giudizi siano basati sul contenuto scientifico delle pubblicazioni, piuttosto che sulle loro metriche di impatto*", allora si dovrebbe fare in modo che questi vengano espressi pubblicamente attraverso la Rete.

I commissari dovrebbero quindi non solo essere competenti e privi di pregiudizi rispetto ai candidati, ma anche molto trasparenti, cioè disposti a condividere l'interessa delle proprie personali riflessioni di merito, avvicinando tutto il processo a una specie di "social peer review".

Il giudizio dei valutatori, potrebbe poi essere a sua volta oggetto di misurazioni statistiche (anch'esse condivise e verificabili dagli osservatori), al pari di quelle dei panelist che assaggiano vini o cibi. Anche il "panel scientometrico" andrebbe, come l'altro, addestrato e opportunamente "tarato".

Così, se un commissario esprimesse un giudizio (statisticamente) troppo distante da quello dei colleghi (outlier), questo andrebbe escluso dalla valutazione complessiva. Avrebbe senso tutto ciò? Forse sì e forse no.

Conclusioni

DORA solleva una questione ineludibile. Perché si possono mettere in campo tutte le più diverse strategie per arrivare ad attribuire il valore scientifico di un determinato prodotto dell'attività di ricerca; sia esso un articolo, un libro, un software, una sequenza genica; alla fine bisognerà comunque arrivare a qualche tipo di classifica e graduatoria sintetica. Questa certamente *distruggerà parte dell'informazione* non potendo tenere, se non parzialmente, conto delle questioni specifiche di merito, ma renderà utilizzabile quella residua, così che detta informazione possa essere utilizzata per attribuire incarichi e valutare un'attività, complessa e multiforme, come quella scientifica.

Nel sistema attualmente in voga, in cui si fa uso di metriche come l'IF, la valutazione è, come giustamente detto da Bertuzzi e Drubin, carente, ma, aggiungiamo noi, tenendo a mente Alessandro Magno, molto "pratica". Il valore dell'IF più basso è zero, quello più alto riportato dal Journal Citation Report (2012) è, al momento attribuito al "CA: a cancer journal for clinicians ([http://onlinelibrary.wiley.com/journal/10.3322/\(ISSN\)1542-4863](http://onlinelibrary.wiley.com/journal/10.3322/(ISSN)1542-4863))", giornale della "American Cancer Society", editore Wiley-Blackwell, che può vantare un IF di 153.459. Chiunque (anche il non esperto), converrebbe che un articolo su CA attribuirebbe allo studioso di oncologia, un impatto (sia pure trattato, normalizzato, percentilizzato in vario modo), davvero alto. Questo tripudio citazionale è ovviamente legato (anche) a fatti che non riguardano solo l'intrinseco valore scientifico degli articoli di CA.

L'IF della rivista in questione è grande perché questa è grandemente letta e citata principalmente perché appartiene a un settore di studi dove:

- (1) ci sono molti progetti di ricerca ben finanziati,
- (2) operano migliaia di ricercatori, per cui
- (3) c'è un'altissima produzione di articoli che generano citazioni.

Inoltre la rivista in questione è molto accessibile in quanto è scritta in inglese ed è localizzata negli States.

Tale preminenza sarà stata ottenuta in tanti modi, ma certamente, e in primo luogo, selezionando, con una revisione paritetica effettuata da capaci scienziati, gli articoli e gli autori. Per cui non si può dire che pubblicare un lavoro su CA sia sostanzialmente equivalente a inserirlo su un'altra qualsiasi rivista sempre in ambito oncologico, ma a impatto quasi 0 (ad esempio, il JCR per il 2012, ne riporta una, recensita dal 2004, che ha un IF di 0.097). Ovviamente nessuno vieta di supporre che articoli di grande valore scientifico, giacciono ignorati in qualche Journal di questo tipo.

"Gi Esami non finiscono mai", è il titolo di una popolare commedia di Eduardo De Filippo (1973), che esprime l'angoscioso stato d'animo, che ognuno ha sperimentato nella sua vita professionale, specialmente se ha operato nella ricerca. In effetti, quando si producono nuovi dati o si propongono interpretazioni dei fatti, si è inevitabilmente sottoposti a giudizio e non sempre in modo benevolo. Dove, però, c'è una procedura, non c'è potere o ce n'è meno. In fondo tutta la questione sollevata dalla Declaration di San Francisco, ruota attorno a questo: al potere, alla prerogativa di valutare gli altri, stabilire i meriti. Cosa pericolosa, specialmente se ci si volesse basare su un dato di natura più soggettiva.

In tal modo, si può facilmente arrivare a classificazioni insidiose che possono assumere risvolti perfino drammatici quando mortificano le aspirazioni personali o ledono gli interessi collettivi.

Prendendo la faccenda con più distacco, queste graduatorie potrebbero apparire perfino

comiche; un po' come nella parodia dell'imprenditore "Al top" di Maurizio Crozza, che per indottrinare giovanissimi, improbabili manager, spara incredibili classifiche su ogni cosa:

*"... portare la nonna al pronto soccorso
sta al 150 posto
dopo la borsa in finta pelle per tenere i documenti in spiaggia..."*

[Citazione a memoria: Maurizio Crozza che imita Flavio Briatore.

Da Crozza nel Paese delle Meraviglie. LA7, 2013]

Per quanto riguarda la nonna siamo certo tutti d'accordo nel rivedere la classifica delle priorità, invece in altri casi tale evidenza non esiste.

Per questo l'iniziativa DORA, pur mobilitando menti e personalità scientifiche di alto livello, probabilmente non riuscirà a stimolare, nel breve periodo, soluzioni di sorta. D'altra parte, la cultura della valutazione, soprattutto per ciò che riguarda la ricerca, non si improvvisa.

Come nella fortunata serie di cartoon della Nickelodeon®, Dora (piccola e avventurosa esploratrice), pone una domanda e rimane in silenzio a guardarci dallo schermo, aspettando la risposta da noi, gli spettatori; che poi, in realtà, saremmo i veri decisori.

Viene da chiedersi: Ma i giovani ricercatori che faranno? Staranno a guardare lo spettacolo o diranno anche loro qualcosa? Si sta parlando proprio di loro.



Dora l'esploratrice. Foto Stephanie (2012).

Creative Commons Attribuzione 2.0 Generico

[copyright vedi: \\storlock\Web Mirror\Web\Provvisorio\corrente\webzine\disc_dora.html]

Riferimenti

Amin M., Mabe M. (2000). Impact Factors: perspectives in publishing: use and abuse. http://cdn.elsevier.com/assets/pdf_file/0014/111425/Perspectives1.pdf.

Ciappelloni R. (2012). Prove di realtà virtuale: affidarsi al Social Peer Review per i decisori istituzionali. *Sanità Pubblica Veterinaria*: Numero 71, Aprile; <http://spvet.it/indice-spv.html#545>.

Eugene Garfield (2006). The history and meaning of the journal impact factor. *JAMA* 295 (1): 90-3.

Figà Talamanca A. (2000). L'Impact Factor nella valutazione della ricerca e nello sviluppo dell'editoria scientifica. <http://siba2.unile.it/sinm/4sinm/interventi/fig-talam.htm>

Garbisa S., Calzà L. (1995). Il peso della qualità accademica ovvero: Manuale per l'uso di parametri oggettivi. Editore Cleup, Padova.

Garfield, Eugene (15 luglio 1955). Citation indexes for science, A new dimension in documentation through association of ideas. *Science* 122 (3159): 108-111. URL consultato in data 18 settembre 2011.

Redazione ROARS (2013). Dichiarazione CUN sulla VQR: "rating e non ranking". <http://www.roars.it/online/dichiarazione-cun-sulla-vqr-rating-e-non-ranking/>, 17 luglio 2013 at 12:42.

Seglen, P.O. (1997). Why the impact factor of journals should not be used for evaluating research. *British Medical Journal*, 314, 497-502.

Stefano Bertuzzi, David G. Drubin (2013). No shortcuts for research assessment. *Molecular Biology of the Cell*, Maggio, vol. 24 no. 10, pag. 1505-1506.



DORA: una critica radicale all'Impact Factor. Come valutare e misurare la qualità dei risultati della ricerca con metodi alternativi? by Ciappelloni R., 2013 is licensed under a Creative Commons Attribuzione - Non commerciale 3.0 Italia License. Permissions beyond the scope of this license may be available at <http://indice.spvet.it/adv.html>.

	Istituto Zooprofilattico Sperimentale dell'Umbria e delle Marche, Via G. Salvemini 1. 06126, Perugia - Italy	
Centralino Istituto	Tel. +39 075 3431 - Fax. +39 075 35047	
Biblioteca	Tel. / Fax +39 075 343217 e-mail: bie@izsum.it	
Rivista SPVet.it ISSN 1592-1581	Tel. +39 075 343207 e-mail: editoria@izsum.it ; redazione-spvet@izsum.it http://spvet.it ; http://indice.spvet.it	
U. R. P.	Tel. +39 075 343223; Fax: +39 075 343289 e-mail: URP@izsum.it	