

MATERIALI

Contro le *unanticipated consequences* della valutazione quantitativa della ricerca: il *Leiden Manifesto* *for research metrics*

di Chiara Faggiolani

A partire dal Pensatore di Rodin

Il Pensatore (*Le Penseur*) (Figura 1) è una delle più celebri sculture della storia dell'arte: un'opera in bronzo realizzata in quasi venticinque anni – tra il 1880 e il 1904 – dallo scultore francese Auguste Rodin e conservata a Parigi nel museo che porta il nome del suo creatore.

Inizialmente il suo nome era 'il Poeta' e faceva parte di una porta monumentale alla quale Rodin lavorò per anni e che avrebbe dovuto essere l'ingresso del Musée des arts decoratifs, mai portato a termine. Per questa porta d'ingresso, Rodin aveva scelto di raffigurare la *Divina Commedia*: il Pensatore doveva rappresentare Dante che, davanti alla porta dell'Inferno, meditava sulla sua grande opera.

Il Pensatore è un uomo nudo, seduto, con il mento appoggiato sul dorso della propria mano destra mentre guarda verso il basso: rappresenta l'essere umano che riflette sul proprio destino e su ciò che lo attende ma anche il pensiero filosofico – come banalmente dimostra una ricerca su Google Immagini con parola chiave 'filosofia'. Filosofia e scienza sono fatte l'una per l'altra, come afferma il fisico John Ziman in un suo bellissimo saggio:

la scienza è un modello di produzione della conoscenza e le sue norme sociali sono inseparabili da quelle epistemiche – definite dai filosofi come principi regolativi della scienza. Non si può sceverare l'idea che gli scienziati hanno su ciò che dovrebbe essere considerata 'la verità' dai modi in cui essi operano insieme per raggiungerla. La filosofia della scienza accademica è, per tanto, parte integrante della sua cultura¹.

Proprio a proposito di scienza, è interessante notare che esistono più di venti esemplari del Pensatore in tutto il mondo e che alcuni sono conservati in importanti università: la Stanford University e la Columbia University ne sono un esempio.

CHIARA FAGGIOLANI, Sapienza Università di Roma, Dipartimento di scienze documentarie, linguistico-filologiche e geografiche, viale Regina Elena 295, 00161 Roma, e-mail chiara.faggiolani@gmail.com.

Ultima consultazione siti web: 14 luglio 2015.

¹ Cfr. John Ziman, *La vera scienza: natura e modelli operativi della prassi scientifica*, Bari: Dedalo, 2002, p. 81 (titolo originale: *Real science: what it is and what it means*, Cambridge: Cambridge University Press, 2000).

Così arriviamo al dunque: lo scorso 23 aprile sulla rivista *Nature* è stato pubblicato *Bibliometrics: Leiden Manifesto for research metrics*², un documento fondamentale per l'accademia, che contiene una serie di raccomandazioni rivolte a rimediare alle ingiustizie introdotte dall'uso scorretto degli indicatori bibliometrici negli esercizi di valutazione della ricerca³.

In apertura al documento una geniale illustrazione di David Parkins⁴ raffigura proprio Il Pensatore di Rodin, la cui testa è stretta dai bracci di un calibro a compasso (Figura 2). Sull'interpretazione della scelta di *Nature* di affidare a questa immagine la rappresentazione del Manifesto torneremo più avanti, dopo averne approfondito il contenuto.

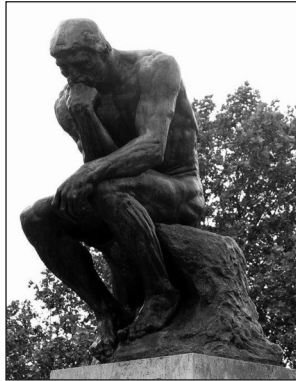


Figura 1 – Il Pensatore di Rodin

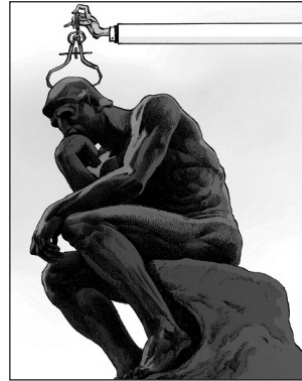


Figura 2 – Illustrazione di David Parkins in apertura al *Bibliometrics: Leiden Manifesto for research metrics*

Il decalogo di Leiden

Nell'ambito del convegno "Context counts: pathways to master big and little data. 19th International conference on Science and technology indicators"⁵, durante una

² Diana Hicks [et al.], *Bibliometrics: the Leiden Manifesto for research metrics*, «Nature», 520 (2015), n. 7548, p. 429-431, <<http://www.nature.com/news/bibliometrics-the-leiden-manifesto-for-research-metrics-1.17351>>.

³ Il Manifesto ha avuto una grande eco, come dimostra il proliferare dei commenti e delle traduzioni in tutte le lingue <<http://www.leidenmanifesto.org/translations.html>>. In Italia si veda il commento uscito su Roars pochi giorni dopo la pubblicazione, <<http://www.roars.it/online/leiden-manifesto-for-research-metrics/>> e Simona Turbanti, *Un Manifesto contro l' 'ossessione della misura'*, «Biblioteche oggi», 33 (2015), n. 6, p. 19-22.

⁴ David Parkins è un noto illustratore inglese che ha disegnato vignette tra gli altri per «The Guardian», «The Observer», «The Sunday Times» e «Nature», <<http://www.davidparkins.com>>.

⁵ Si tratta di una delle principali occasioni di incontro annuale per gli specialisti di metriche per la valutazione della scienza e della tecnologia <<http://stiz014.cwts.nl/Home>>, organizzato dal prestigioso Centre for Science and Technology Studies (CWTS) dell'Università di Leiden. Il CWTS si occupa di metodi bibliometrici avanzati, pratiche di valutazione nelle scienze umane e sociali, impatto sociale della ricerca ecc.; è una delle principali scuole bibliometriche a livello internazionale, <www.cwts.nl>.

sessione dedicata agli standard di qualità per gli indicatori in scienza e tecnologia⁶, la comunità degli scientometristi, forse ritenendo fosse arrivato il momento di avviare una sorta di opera moralizzatrice, ha riconosciuto in modo pubblico, chiaro e partecipato le cosiddette *unanticipated consequences*⁷, le conseguenze inattese o, potremmo anche dire, gli effetti indesiderati, derivanti dall'uso poco accorto degli indicatori bibliometrici nella valutazione della ricerca scientifica.

È stato quindi formulato un documento che, come anticipato, è stato pubblicato ad aprile 2015 sulla prestigiosa rivista *Nature* a firma di cinque studiosi: Diana Hicks del Georgia Institute of Technology, Paul Wouters, direttore del Centre for Science and Technology Studies (CWTS) dell'Università di Leiden, Ludo Waltman e Sarah de Rijcke – sempre del gruppo di Leiden – e Ismael Rafols dell'Instituto de Gestión de la Innovación y del Conocimiento (Ingenio)⁸ del Consejo Superior de Investigaciones Científicas (CSIC) dell'Università Politecnica di Valencia (UPV).

Bibliometrics: Leiden Manifesto for research metrics contiene dieci principi fondamentali (Figura 3) che chiunque utilizzi gli indicatori bibliometrici a scopo valutativo deve applicare, pena la possibilità che «il sistema venga danneggiato da quegli stessi strumenti pensati e progettati per migliorarlo»⁹.

-
- 1 *La valutazione quantitativa deve essere di supporto alla valutazione qualitativa dei pari*

 - 2 *Misurare le prestazioni in relazione agli obiettivi della ricerca dell'istituzione, del gruppo o del ricercatore*

 - 3 *Salvaguardare l'eccellenza della ricerca a livello locale*

 - 4 *Il processo di raccolta e analisi dei dati deve essere aperto, trasparente e semplice*

 - 5 *Consentire ai valutati di verificare i dati e l'analisi utilizzati nella valutazione*

 - 6 *Tenere in considerazione la diversità dei comportamenti citazionali nelle pratiche di pubblicazione nelle diverse discipline scientifiche*

 - 7 *La valutazione dei singoli ricercatori deve basarsi su un giudizio qualitativo del loro portfolio di ricerca*

 - 8 *Evitare finta concretezza e falsa precisione*

 - 9 *Si devono riconoscere gli effetti sistemici della valutazione e degli indicatori*

 - 10 *Gli indicatori devono essere verificati e aggiornati regolarmente*

Figura 3 – Il decalogo di Leiden

Di seguito si propone una traduzione non perfettamente letterale ma esaustiva dei dieci punti.

6 Per un puntuale resoconto della sessione si veda il blog «The citation culture» di Paul Wouters e Sarah de Rijcke, September 15, 2014, <<https://citationculture.wordpress.com/2014/09/15/the-leiden-manifesto-in-the-making-proposal-of-a-set-of-principles-on-the-use-of-assessment-metrics-in-the-st-indicators-conference/>>.

7 Cfr. Robert K. Merton, *The unanticipated consequences of purposive social action*, «American sociological review», 1 (1936), p. 894-904, <<http://www.d.umn.edu/cla/faculty/jhamlin/4111/Readings/MertonSocialAction.pdf>>.

8 Ingenio, <<http://www.ingenio.upv.es/en>>.

9 D. Hicks [et al.], *Bibliometrics: the Leiden Manifesto for research metrics* cit., p. 429.

1. *La valutazione quantitativa deve essere di supporto alla valutazione qualitativa dei pari*
 Gli indicatori possono ridurre la tendenza al pregiudizio insita nella *peer review* e facilitare il giudizio. In questo senso, gli indicatori possono rafforzare la valutazione dei pari, posto che dare giudizi sui colleghi è difficile senza una serie di informazioni rilevanti. Tuttavia, i valutatori non devono lasciarsi tentare dalla possibilità di delegare il processo decisionale ai numeri. Gli indicatori non devono sostituirsi a un giudizio consapevole.

I valutatori devono mantenere piena responsabilità dei propri giudizi.

2. *Misurare le prestazioni in relazione agli obiettivi della ricerca dell'istituzione, del gruppo o del ricercatore*

Gli obiettivi di un programma di ricerca dovrebbero essere indicati sin dall'inizio e gli indicatori utilizzati per valutarne la performance devono riferirsi chiaramente a questi obiettivi. La scelta degli indicatori e il modo in cui vengono utilizzati devono tenere in considerazione il contesto socio-economico e culturale. Gli scienziati hanno obiettivi di ricerca diversi. La ricerca finalizzata all'avanzamento della conoscenza è molto diversa dalla quella finalizzata ad offrire risposte ai problemi della società. La valutazione può essere basata sul riconoscimento di meriti che sono rilevanti per la politica, per l'industria o per i cittadini in generale, piuttosto che sull'idea accademica di eccellenza scientifica.

Non esiste un unico modello di valutazione che sia applicabile in tutti i contesti.

3. *Salvaguardare l'eccellenza della ricerca a livello locale*

In molte parti del mondo, l'eccellenza della ricerca è equiparata alla pubblicazione in lingua inglese. La legge spagnola, per esempio, afferma la volontà degli studiosi spagnoli di pubblicare in riviste ad alto impatto. L'*impact factor* (IF) è calcolato automaticamente per le riviste indicizzate in Web of science, un database che indicizza per la gran parte riviste americane in lingua inglese. Questi pregiudizi sono particolarmente problematici nelle scienze umane e sociali, aree in cui la ricerca si concentra in tematiche nazionali o locali. Ma non solo: molti altri campi di ricerca hanno una dimensione nazionale o regionale, per esempio in ambito epidemiologico lo studio dell'HIV nell'Africa sub-sahariana.

Il pluralismo e la rilevanza sociale tendono ad essere ignorati nel produrre articoli di interesse per i sostenitori dell'alto impatto: le riviste in lingua inglese. I sociologi spagnoli molto citati in Web of science hanno lavorato su modelli astratti o su studi basati su dati americani. Così facendo si è persa la specificità dei sociologi che hanno un alto impatto nelle riviste in lingua spagnola: argomenti come il diritto del lavoro locale, l'assistenza sanitaria per gli anziani o l'occupazione per gli immigrati¹⁰. Per identificare e premiare l'eccellenza della ricerca rilevante a livello locale sarebbero necessarie metriche costruite per la letteratura scientifica di alta qualità non in lingua inglese.

4. *Il processo di raccolta e analisi dei dati deve essere aperto, trasparente e semplice*

La costruzione dei database necessari per la valutazione deve seguire processi chiaramente indicati prima che la ricerca sia stata completata. Questa era una pratica comune tra i diversi gruppi accademici e commerciali che hanno costruito metodologie di valutazione bibliometrica nei decenni passati: tali gruppi pubblicavano i protocolli di riferimento nelle riviste sottoposte a *peer review*. Questa trasparenza ha consentito un controllo scrupoloso del metodo. Ad esempio, nel 2010, il dibattito pubblico sulle caratteristiche tecniche di un importante indicatore utilizzato da uno dei nostri gruppi (il Centre for Science

10 Carla López Piñeiro; Diana Hicks, *Reception of Spanish sociology by domestic and foreign audiences differs and has consequences for evaluation*, «Research Evaluation», 24 (2015), n. 1, p. 78-89.

and Technology Studies, CWTS, dell'Università di Leiden nei Paesi Bassi) ha portato a una revisione nel calcolo di questo indicatore¹¹. Gli operatori commerciali nuovi nel settore dovrebbero attenersi agli stessi standard. Nessuno dovrebbe accettare una valutazione a 'scatola nera'. La semplicità è una virtù per un indicatore perché ne aumenta la trasparenza. Ma metriche semplicistiche possono distorcere la valutazione (vedi principio 7).

I valutatori devono sforzarsi di bilanciare la semplicità degli indicatori con la complessità del processo di ricerca.

5. *Consentire ai valutati di verificare i dati e l'analisi utilizzati nella valutazione*

Per garantire la qualità dei dati, a tutti i ricercatori inclusi negli studi bibliometrici deve essere data la possibilità di controllare che i loro prodotti siano stati identificati correttamente. Tutti coloro che dirigono e gestiscono i processi di valutazione devono assicurare l'accuratezza dei dati, attraverso l'autoverifica o audit da parte di terzi. Le università potrebbero attuare questo meccanismo nei loro repository. Questo dovrebbe essere un principio guida nella selezione dei fornitori di questi sistemi. Raccogliere ed elaborare dati precisi e di alta qualità richiede tempo e denaro. I responsabili devono assegnare risorse che siano adeguate a questo livello di qualità.

6. *Tenere in considerazione la diversità dei comportamenti citazionali nelle pratiche di pubblicazione nelle diverse discipline scientifiche*

La migliore pratica di valutazione è selezionare una batteria di possibili indicatori e consentire alle diverse comunità scientifiche di scegliere quello che più le rappresenta. Alcuni anni fa, un gruppo di storici europei ha ricevuto un punteggio relativamente basso in una valutazione nazionale con *peer review* perché ha scritto libri piuttosto che articoli su riviste indicizzate in Web of science. Questi storici hanno avuto la sfortuna di far parte di un dipartimento di psicologia. La valutazione di storici e di scienziati sociali necessita l'inclusione di libri e della letteratura nazionale tra le pubblicazioni che contano, così come per gli informatici vanno presi in considerazione gli atti di convegno. I *range* citazionali variano da disciplina a disciplina: le riviste *top* in matematica hanno un IF circa di 3; i *top journals* in biologia hanno un IF circa di 30.

Per questo sono necessari indicatori normalizzati, e il metodo di normalizzazione più robusto è basato sui percentili: ogni articolo è ponderato sulla base del percentile di appartenenza nella distribuzione citazionale del suo campo (per esempio il percentile 1%, 10% o 20%). Una singola pubblicazione altamente citata migliora leggermente la posizione di una università in un ranking basato sui percentili, ma può spingere l'università dal centro verso la cima di una classifica costruita sulla media delle citazioni¹².

7. *La valutazione dei singoli ricercatori deve basarsi su un giudizio qualitativo del loro portfolio di ricerca*

Più si è anziani più alto sarà l'indice H^{13} , anche in assenza di nuove pubblicazioni. L'indice H varia a seconda del settore: gli scienziati top delle scienze della vita pos-

11 Anthony F. J. van Raan [et al.], *Rivals for the crown: reply to Opthof and Leydesdorff*, «Journal of Informetrics», 4 (2010), n. 3, p. 431-435, <<http://www.cwts.nl/TvR/documents/AvR-2010JolRivals.pdf>>.

12 Ludo Waltman [et al.], *The Leiden ranking 2011/2012: data collection, indicators, and interpretation*, «Journal of the American Society for Information Science and Technology», 63 (2012), n. 12, p. 2419-2432.

13 Indice H, o indice di Hirsch (H-index) è un indicatore presentato 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 e sul numero di citazioni ricevute (n.d.R.).

sono arrivare a 200; i fisici a 100 e gli scienziati sociali a 20-30¹⁴. L'indice H dipende dal database: ci sono ricercatori in informatica che hanno un indice H di circa 10 calcolato a partire da Web of science, ma di 20-30 in Google Scholar¹⁵.

Leggere e giudicare il lavoro di uno studioso è molto più appropriato che basarsi su un unico numero. Anche quando si confrontino grandi numeri di ricercatori, è meglio adottare un approccio capace di considerare diverse informazioni sulle competenze, le esperienze, le attività, il prestigio.

8. Evitare finta concretezza e falsa precisione

Gli indicatori sono inclini all'ambiguità concettuale e all'incertezza e richiedono forti assunzioni che non sono universalmente accettate. Il significato dei conteggi citazionali, per esempio, è stato a lungo dibattuto. Per queste ragioni, le migliori pratiche utilizzano più indicatori per fornire un quadro più robusto e pluralista. Se incertezza ed errore possono essere quantificati, ad esempio utilizzando le *error bars*, queste informazioni devono accompagnare i valori degli indicatori pubblici. Se ciò non è possibile, i produttori degli indicatori dovrebbe almeno evitare falsa precisione. Ad esempio, l'IF è pubblicato con tre cifre decimali per evitare pareggi. Tuttavia, data l'ambiguità concettuale e la variabilità casuale dei conteggi citazionali, non ha senso distinguere tra riviste sulla base di differenze minime dell'IF. Si deve evitare falsa precisione: è giustificato l'uso di un solo decimale.

9. Si devono riconoscere gli effetti sistemici della valutazione e degli indicatori

Gli indicatori cambiano il sistema scientifico attraverso gli incentivi che stabiliscono. Questi effetti dovrebbero essere previsti. Ciò significa che è sempre meglio una batteria di indicatori dato che uno solo inviterà a comportamenti opportunistici e a uno spostamento degli obiettivi (in cui la misurazione diviene lo scopo). Ad esempio negli anni Novanta, l'Australia ha finanziato la ricerca universitaria con una formula basata in gran parte sul numero di articoli pubblicati dagli istituti. Le università potevano calcolare il 'valore' di un articolo pubblicato in una rivista referata; nel 2000 il valore si stimò in 800 dollari australiani (circa 480 dollari US nel 2000) in finanziamenti per la ricerca. Com'era prevedibile, il numero di articoli pubblicati dai ricercatori australiani è aumentato ma in riviste meno citate, evidenziando un crollo della qualità della ricerca¹⁶.

10. Gli indicatori devono essere verificati e aggiornati regolarmente

Le missioni di ricerca, gli obiettivi della valutazione e lo stesso sistema di ricerca cambiano. Le metriche che una volta erano utili diventano inadeguate e ne nascono delle altre. Per questa ragione, i sistemi di indicatori devono essere verificati e se necessario modificati. Rendendosi conto degli effetti della sua formula semplicistica, l'Australia nel 2010 ha adottato il sistema ERA, *Excellence in Research for Australia*, che attribuisce maggiore valore alla qualità.

14 Jorge E. Hirsch, *An index to quantify an individual's scientific research output*, «Proceedings of the National Academy of Sciences of the United States of America», 102 (2005), n. 46, p. 16569-16572, <<http://arxiv.org/abs/physics/0508025>>.

15 Judit Bar-Ilan, *Which h-index? A comparison of WoS, Scopus and Google Scholar*, «Scientometrics», 74 (2008), n. 2, p. 257-271.

16 Linda Butler, *Explaining Australia's increased share of ISI publications: the effects of a funding formula based on publication counts*, «Research policy», 32 (2003), n. 1, p. 143-155.

Il Manifesto conclude evidenziando i passaggi successivi: rispettando questi dieci principi – dicono gli autori – la valutazione della ricerca può giocare un importante ruolo nello sviluppo della scienza e del suo rapporto con la società: quest'ultimo aspetto sempre più fondamentale – mi permetto di aggiungere.

Gli indicatori possono fornire informazioni importanti che sarebbe difficile recuperare a partire soltanto dal giudizio degli esperti. Allo stesso tempo, però, non si deve permettere che la valutazione quantitativa si converta essa stessa in un obiettivo. Le decisioni migliori si prendono facendo riferimento sì a statistiche robuste, ma che siano sensibili rispetto agli obiettivi e alla natura della ricerca che viene valutata. È dunque necessario tanto un approccio quantitativo quanto un approccio qualitativo. Le decisioni che riguardano la scienza devono essere sempre prese in base a processi di alta qualità, basati su dati anch'essi di alta qualità.

La portata del Manifesto: da DORA a Leiden

I principi di questo Manifesto non sono nuovi nei contenuti¹⁷.

Già nel 2012 nella *San Francisco Declaration of research assessment-DORA*¹⁸ – documento redatto dall'American Society for Cell Biology – erano stati raccolti diciotto principi specificatamente rivolti ai cinque operatori che entrano nel processo di valutazione della ricerca: agenzie di finanziamento, istituzioni, editori, fornitori di metriche per la valutazione e ricercatori (Figura 4).

RACCOMANDAZIONE GENERALE	1	<i>Non utilizzare gli indicatori bibliometrici (IF ad esempio) come misura sostitutiva della qualità della ricerca</i>
RACCOMANDAZIONI PER LE AGENZIE DI FINANZIAMENTO	2	<i>Rendere espliciti i criteri usati per valutare la produttività scientifica e sottolineare chiaramente che il contenuto scientifico di un articolo è molto più importante delle metriche della pubblicazione e dell'identità della rivista nel quale è pubblicato.</i>
	3	<i>Considerare il valore e l'impatto di tutti i prodotti della ricerca (inclusi i datasets e il software) in aggiunta alle pubblicazioni, e considerare un ampio insieme di misure di impatto che includano indicatori qualitativi dell'impatto della ricerca, come l'influenza sulle politiche e sulle pratiche</i>

¹⁷ Il 12 maggio 2014 a Parigi, nell'ambito di un workshop internazionale dedicato a questi temi, si era già discusso di linee guida per una corretta applicazione degli indicatori bibliometrici, <http://www.obs-ost.fr/fractivités/workshop_international>. Gli stessi autori del Manifesto affermano «Its ten principles are not news to scientometricians, although none of us would be able to recite them in their entirety because codification has been lacking until now». Cfr. D. Hicks [et al.], *Bibliometrics: the Leiden Manifesto for research metrics* cit., p. 429.

¹⁸ San Francisco Declaration on research assessment (DORA), *San Francisco Declaration on research assessment. Putting science into the assessment of research*, <<http://www.ascb.org/dora-old/files/SFDeclarationFINAL.pdf>>.

RACCOMANDAZIONI PER LE ISTITUZIONI	4	<i>Vedi principio 2</i>
	5	<i>Vedi principio 3</i>
RACCOMANDAZIONI PER GLI EDITORI	6	<i>Ridurre l'enfasi sull'IF come strumento di promozione delle riviste, offrendo una pluralità di metriche ed informazioni in grado di fornire un quadro più ricco sulla performance della singola rivista</i>
	7	<i>Rendere disponibili metriche relative ai singoli articoli pubblicati</i>
	8	<i>Incoraggiare pratiche responsabili per l'individuazione degli autori e per fornire informazioni sullo specifico contributo di ciascun autore</i>
	9	<i>Adottare la Creative Commons Public Domain Dedication alle riviste open-access</i>
	10	<i>Eliminare limiti al numero di riferimenti bibliografici negli articoli scientifici e, quando appropriato, favorire la citazione dei lavori di ricerca originali anziché il rimando a rassegne</i>
RACCOMANDAZIONI PER LE ORGANIZZAZIONI CHE FORNISCONO METRICHE	11	<i>Adottare la trasparenza sui dati e sui metodi usati per calcolare tutte le metriche</i>
	12	<i>Fornire i dati con licenze che non ne limitino il riuso</i>
	13	<i>Definire regole per combattere la manipolazione degli indicatori</i>
	14	<i>Definire indicatori precisi in riferimento a differenti campi di indagine e tipologie di prodotti di ricerca (articoli e rassegne)</i>
RACCOMANDAZIONI PER I RICERCATORI	15	<i>Quando coinvolti in comitati che prendono decisioni sui finanziamenti, assunzioni o promozioni, fare valutazioni in base al contenuto scientifico, piuttosto che alle metriche di pubblicazione.</i>
	16	<i>Citare la letteratura primaria nella quale le osservazioni sono documentate per la prima volta invece di articoli di rassegna al fine di attribuire credito a chi lo merita</i>

RACCOMANDAZIONI PER I RICERCATORI	17	<i>Usare un insieme di metriche e indicatori per gli articoli [...] come evidenza dell'impatto dei singoli articoli pubblicati e degli altri prodotti della ricerca</i>
	18	<i>Mettere in discussione le pratiche di valutazione della ricerca che si basano impropriamente sui fattori di impatto delle riviste e promuovere le buone pratiche che sottolineano il valore e l'influenza dei singoli prodotti della ricerca</i>

Figura 4 – I diciotto principi di DORA

Sulla necessità dell'applicazione di una seria *informed peer review* negli esercizi di valutazione della ricerca (principio 1) sembra esserci ormai una totale convergenza: come dimostrano anche alcune tra le principali esperienze internazionali – si pensi al britannico REF *Research Excellence Framework* – e alcuni documenti elaborati dalle principali istituzioni europee sulla regolamentazione della *peer review*¹⁹.

La novità del Manifesto sta piuttosto nella forma in cui i contenuti sono espressi. Due sono essenzialmente i punti di forza che consentono di considerarlo una 'pietra miliare' nella strada della valutazione della ricerca.

Il primo punto di forza dei dieci principi di Leiden sta nella loro enunciazione generale. Sono i bibliometristi stessi a mettere in discussione l'efficacia degli indicatori bibliometrici nella valutazione della ricerca, non soltanto nelle scienze umane e sociali ma, in generale, anche rispetto alle scienze dure. Basta osservare chi sono i promotori del Manifesto per capirne la portata. Leggendo il decalogo non stupisce che tra gli autori figurino Diana Hicks, studiosa molto nota a chi si occupa di valutazione della ricerca nelle scienze umane e sociali (SSH) che, con il celebre saggio sulle 'quattro letterature'²⁰ del 2004, aveva già messo in evidenza come il modello pro-

¹⁹ Anche le linee guida sulla *peer review*, elaborate nel 2011 dalla European Science Foundation, mettono in evidenza chiaramente come gli indicatori bibliometrici debbano essere utilizzati con cautela solo come strumento complementare (*complementary tool*) nella valutazione dei pari. Cfr. European Science Foundation, *European peer review guide. integrating policies and practices into coherent procedures*, 2011, <http://www.esf.org/fileadmin/Public_documents/Publications/European_Peer_Review_Guide_01.pdf>. Si veda anche Institut de France, Academie des sciences, *Du bon usage de la bibliometrie pour l'evaluation individuelle des chercheurs*, 17 January 2011, <<http://www.academie-sciences.fr/fr/Rapports-ouvrages-avis-et-recommandations-de-l-Academie/du-bon-usage-de-la-bibliometrie-pour-l-evaluation-individuelle-des-chercheurs.html>>.

²⁰ Cfr. Diane Hicks, *The four literatures for social science*. In *Handbook of quantitative science and technology research*, a cura di Henk F. Moed, Wolfgang Glänzel, Ulrich Schmoch. Dordrecht: Kluwer Academic Publisher, 2004, p. 473-496. In questo saggio l'autrice individua quattro tipi di pubblicazioni per le SSH: gli articoli su riviste internazionali, le monografie e tutto ciò che può essere riferito alla forma libro, le riviste di carattere nazionale e la letteratura divulgativa. I database commerciali Web of science e Scopus, coprono in maniera davvero modesta i libri e le riviste nazionali e non coprono affatto la letteratura di 'divulgazione'.

posto da Web of science o Scopus²¹ non rispecchi (e non rispetti) le modalità di disseminazione della conoscenza e dei risultati della ricerca delle SSH. È, invece, decisamente più sorprendente osservare tra gli autori la presenza di Paul Wouters e degli altri studiosi del CWTS di Leiden, cioè il centro che «elabora e che produce e vende report bibliometrici sulla base dei quali istituzioni e nazioni orientano le loro politiche della ricerca»²².

Come sottolineano i principi 3 e 6 del Manifesto, se andiamo nello specifico delle diverse discipline, anche in ambito *hard science*, osserviamo comportamenti citazionali e modalità di disseminazione dei risultati della ricerca molto diversi.

In questo senso, facendo riferimento all'esperienza italiana, uno dei principali lasciti della VQR 2004-2010²³, a mio avviso, è proprio quello di aver messo nero su bianco tale aspetto. Leggendo i quattordici rapporti d'area dei GEV dell'ANVUR²⁴ si comprende, infatti, come in un'ottica valutativa esistano specificità non solo a livello di area o macro-settore ma anche all'interno dei singoli settori scientifico-disciplinari, specificità che richiedono scelte metodologiche del tutto diverse. Tutti gli studiosi, soprattutto i più giovani, dovrebbero leggere il rapporto GEV della propria area di appartenenza per capire meglio in quali acque stanno navigando.

Il secondo punto di forza – conseguenza del primo – sta nell'aver sancito in modo inequivocabile il legame tra l'uso distorto degli indicatori bibliometrici utilizzati a scopo valutativo e le *unanticipated consequences* che possono derivare nello sviluppo della scienza (principi 3, 6, 8, 9). Quello di *unanticipated consequences*, come già detto, è un concetto introdotto per la prima volta dal sociologo della scienza Robert K. Merton in un suo articolo del 1936 sull'*American sociological review*²⁵. Si tratta delle conseguenze inattese: se gli indicatori bibliometrici cessano di essere una misura e diventano un obiettivo e dunque, se la citazione diventa una moneta di scambio, non è difficile immaginare che, invece di pensare a pubblicare buona ricerca, i ricercatori inizierebbero a preoccuparsi di pubblicare ricerca meno buona ma su riviste con elevato IF, così come potrebbe esserci il rischio di un appiattimento su tematiche di ricerca *mainstream*, poiché più citate, a scapito di una sana pluralità tematica e metodologica.

21 Per una panoramica generale sui principali database bibliometrici e sugli indicatori mi permetto di rinviare a Chiara Faggiolani, *La bibliometria*, Roma: Carocci, 2015.

22 Roars, *Leiden Manifesto for research metrics*, <<http://www.roars.it/online/leiden-manifesto-for-research-metrics/>>.

23 Nel 2010 è stata istituita in Italia l'Agenzia nazionale di valutazione del sistema universitario e della ricerca (ANVUR), che ha gestito il secondo esercizio nazionale di valutazione della qualità della ricerca (VQR) per il periodo 2004-2010, il più grande realizzato finora. La VQR 2004-2010 ha coinvolto obbligatoriamente università statali ed enti di ricerca vigilati dal Miur, più altre strutture che hanno chiesto volontariamente di essere sottoposte alla valutazione. Il metodo di valutazione utilizzato è la cosiddetta *informed peer review*: per ciascuna delle 14 aree disciplinari CUN è stato nominato un GEV (gruppo di esperti di valutazione), formato da studiosi di nota esperienza e valenza scientifica che hanno avuto il compito di gestire la valutazione delle pubblicazioni (3 per ogni studioso), affiancando alla *peer review* – ove possibile – gli indicatori bibliometrici per formulare il loro giudizio sui soggetti sottoposti a valutazione.

24 ANVUR, *Valutazione della Qualità della Ricerca 2004-2010 (VQR 2004-2010). Rapporto finale*, 30 giugno 2013, <<http://www.anvur.org/rapporto/>>. I rapporti sono scaricabili con le relative tabelle.

25 Cfr. R.K. Merton, *The Unanticipated Consequences of Purposive Social Action* cit. p. 894-904.

Questi strumenti, se male utilizzati, possono essere molto pericolosi. Gli indicatori bibliometrici, utilizzati a scopo valutativo, portano con sé i pregiudizi dei database soggiacenti²⁶. Questa considerazione è trasversale alle diverse discipline scientifiche: il limite delle metriche quantitative non deriva perciò soltanto dalla carenza di dati, come tanto si è detto a proposito delle scienze umane e sociali, ma anche dalla loro naturale impossibilità di intercettare e interpretare la qualità, tanto delle pubblicazioni quanto degli studiosi.

Conclusioni

Una seria riflessione metodologica – sul come – non può non tenere in considerazione l'oggetto della valutazione – il cosa – e, allo stesso modo, non può tralasciare il contesto in cui essa si concretizza e il suo obiettivo – il perché. Entrambi, infatti, concorrono alla definizione dell'apparato metodologico che risulta essere il più adeguato.

Affinché la valutazione della ricerca fornisca informazioni necessarie al miglioramento della ricerca stessa – il fine ultimo di qualsiasi attività di valutazione – gli strumenti devono essere pensati non in modo asettico ma devono rivolgere attenzione al contesto: ovvero, alla natura delle sfide con cui la ricerca è chiamata a confrontarsi oggi e in futuro, e alle risorse attraverso le quali essa si può concretizzare.

Si sta assistendo a una radicale trasformazione del modo in cui la scienza è organizzata, gestita e resa operativa²⁷, anche gli strumenti per la valutazione della ricerca devono essere ripensati, pena le conseguenze inattese di cui si è detto.

In conclusione, tornando all'illustrazione di David Parkins e senza dilungarci troppo nell'esercizio iconologico, focalizziamo l'attenzione su tre aspetti:

- 1) la testa del Pensatore è stretta dai bracci di un calibro a compasso (per esterni): uno strumento adatto a misurare la distanza tra due facce di una concavità, la larghezza di un oggetto, ma ovviamente non la sua materia, la sua densità, il suo contenuto, la sua qualità;
- 2) la precisione di questo strumento dipende fortemente dalla perizia con la quale viene utilizzato: può fornire misurazioni maldestre oppure stime precise al centesimo di millimetro. Dipende da diversi fattori;
- 3) il calibro è trattenuto dalla punta delle dita nervose di una mano in cui culmina un rigido e lungo braccio. Chi misura è distante, non può rendersi conto della puntualità della sua misurazione.

Mi fermo qui perché credo sia evidente, rispetto alle nostre finalità e per quanto detto, che quello scelto è uno strumento inadeguato in relazione all'obiettivo che metaforicamente la misurazione del cranio rappresenta: misurare la scienza, il talento, l'arte di uno studioso non è cosa da calibro a compasso, non è cosa da misurazione quantitativa.

²⁶ In letteratura è noto il caso di Ike Antkare, studioso inventato di sana pianta da un gruppo di ricercatori francesi per dimostrare la mancanza di affidabilità di una valutazione basata su criteri strettamente bibliometrici. Ike Antkare, con 102 articoli (tutti scritti nel 2009), ha ottenuto un indice H di 94. Per questo si colloca al 21° posto della classifica generale di GS degli scienziati più citati: meno di Freud che è al primo posto con un indice H di 183 ma meglio di Einstein, che ha un indice H di 84 e si colloca al 36° posto. Si rimanda all'articolo stesso per i dettagli metodologici. Cfr. Cyril Labbè, *Ike Antkare one of the great stars in the scientific firmament*, «International Society for Scientometrics and Informetrics Newsletter», 6 (2010), n. 2, p. 48-52, <<https://hal.inria.fr/hal-00713564/document>>.

²⁷ Il passaggio dalla scienza accademica alla scienza post-accademica è stato approfondito da Chiara Faggiolani, *Valutazione della ricerca e scienza mode 2*, «Lettera matematica Pristem», (2013), n. 87, p. 34-40.

Il 'Pensatore misurato', in apertura al Manifesto, nella mia interpretazione²⁸, esprime proprio questo: il rischio di scelte valutative sbagliate e la conseguente possibilità di una rottura degli equilibri interni grazie ai quali la scienza si sviluppa: «[...] la scienza è di per sé un sistema sociale complesso, auto-organizzato, che si adatta in maniera opportunistica al mutare delle circostanze»²⁹.

Articolo proposto il 29 giugno 2015 e accettato il 30 giugno 2015.

ABSTRACT AIB studi, vol. 55 n. 3 (settembre/dicembre 2015), p. 427-438. DOI 10.2426/aibstudi-11366

CHIARA FAGGIOLANI, Sapienza Università di Roma, Dipartimento di scienze documentarie, linguistico-filologiche e geografiche, viale Regina Elena 295, 00161 Roma, e-mail chiara.faggiolani@gmail.com.

Contro le *unanticipated consequences* della valutazione quantitativa della ricerca: il *Leiden Manifesto for research metrics*

Durante il convegno "STI-Science and technology indicators 2014", una delle principali occasioni di incontro annuale per bibliometristi e scientometristi – organizzato dal prestigioso Centre for Science and Technology Studies (CWTS) dell'Università di Leiden – è nato un documento che riconosce in modo pubblico, chiaro e partecipato le cosiddette *unanticipated consequences* (conseguenze inattese) derivanti da un uso non accorto degli indicatori bibliometrici nella valutazione della ricerca. Il *Bibliometrics: Leiden Manifesto for research metrics*, pubblicato lo scorso aprile su «Nature», fissa i 10 principi fondamentali rivolti a chi utilizza le metriche quantitative per scopi valutativi.

L'articolo propone un commento a margine di questo documento: partendo dalle differenze rispetto alla precedente *San Francisco Declaration of research assessment-DORA*, vengono messi in evidenza i punti di forza del manifesto, che ci permettono di considerarlo una pietra miliare sulla strada della valutazione della ricerca.

Against the *unanticipated consequences* of quantitative research evaluation: the *Leiden Manifesto for research metrics*

During the annual conference "STI-Science and technology indicators 2014", one of the main opportunities for specialists of science and technology metrics to meet each other – organized by the prestigious Centre for Science and Technology Studies (CWTS) at the University of Leiden – a group of researchers produced a document that clearly recognizes the *unanticipated consequences* resulting from the pervasive misapplication of indicators to the evaluation of scientific performance.

The article *Bibliometrics: Leiden Manifesto for research metrics*, published on April, 2015 on «Nature», establishes the 10 fundamental principles for anyone using quantitative metrics for evaluation purposes. Stressing the differences between the previous *San Francisco Declaration of research assessment-DORA* and the *Manifesto*, the paper emphasizes the strengths of the latter, that can be rightly considered a milestone on the road to evaluation research.

28 L'iconologia di un'immagine, finalizzata a ricostruire il contesto in cui è nata attraverso la ricerca dei significati simbolici nascosti, porta sempre con sé il rischio di sovrainterpretazioni o, nei casi peggiori, di interpretazioni sbagliate. L'interpretazione del 'Pensatore misurato', proposta in queste pagine, è il frutto di uno scambio di riflessioni con Giovanni Solimine, che desidero ringraziare, mantenendo comunque totale responsabilità per la visione suggerita.

29 Cfr. J. Ziman, *La vera scienza: natura e modelli operativi della prassi scientifica* cit., p. 96.