

**Un *caveat* per apprendisti stregoni bibliometrici.
Ancora a proposito di bibliometria, scienze umane e procedure
di valutazione in Italia**

di Antonio Banfi

During the Middle Ages there were all kinds of crazy ideas, such as that a piece of rhinoceros horn would increase potency. Then a method was discovered for separating the ideas--which was to try one to see if it worked, and if it didn't work, to eliminate it. This method became organized, of course, into science. And it developed very well, so that we are now in the scientific age. It is such a scientific age, in fact, that we have difficulty in understanding how witch doctors could ever have existed, when nothing that they proposed ever really worked--or very little of it did.

R. Feynman, Cargo Cult Science (1974)

Abstract: in the following pages, we attempt a preliminary analysis of the new bibliometrical project for the social and human sciences which will be launched on January the 20th by the Italian National Agency for the evaluation of higher education and research. Although evaluation systems that are not based on peer review only are recommended for scientific fields such as these, however the project, at a first glance, has a number of flaws and could therefore adversely affect the development of Italian research in the SSH. Further reflection and discussion is required: citation analysis cannot be the only bibliometric tool available for the evaluation of SSH. More caution, more awareness of the international debate and, ultimately, the development of more sophisticated indicators are necessary.

Il 20 gennaio 2014 si terrà a Roma un incontro, promosso dall'Agenzia nazionale per la valutazione della ricerca e del sistema universitario (ANVUR), incentrato sul tema assai complesso della valutazione delle scienze umane (d'ora in avanti SSH).¹ Nel corso di tale incontro l'Agenzia annuncerà alla comunità scientifica i propri progetti relativi alla costruzione di un database bibliometrico (citazionale e bibliografico) da utilizzarsi

¹ La documentazione relativa all'incontro è disponibile all'indirizzo http://www.anvur.org/index.php?option=com_content&view=article&id=570&Itemid=545&lang=it

proprio a questo scopo. Nelle pagine che seguono desidero richiamare l'attenzione dei lettori sugli aspetti critici di tale iniziativa, nella speranza che sia l'Agenzia, sia le comunità scientifiche, prendano coscienza dei rischi derivanti da alcune scelte preannunciate nella documentazione resa finora disponibile, e che si voglia per una volta procedere con metodo scientifico nell'esplorazione di questi nuovi territori: essi sono noti solo parzialmente, e non è allo stato disponibile una documentazione né una ricerca scientometrica matura. Mi permetto dunque di avanzare un sommesso richiamo alla prudenza e al senso di responsabilità, perché si eviti di far danno.

La breve analisi che segue è articolata in tre parti. La prima attiene all'architettura del sistema di valutazione; la seconda ai problemi posti dall'uso della bibliometria nelle SSH; la terza concerne scelte potenzialmente pericolose per l'integrità (etica e scientifica) della ricerca nelle SSH. Infine concluderò con alcuni suggerimenti.

Prima di tutto ciò, è opportuno porsi una domanda: perché l'analisi bibliometrica delle SSH?

La risposta è tutto sommato semplice: la revisione dei pari, da molti considerata come l'unico strumento sicuro ai fini della valutazione di queste aree disciplinari nel loro complesso,² è estremamente costosa, sia in termini di tempo che di impegno delle risorse umane. E' noto che il REF inglese fa ricorso in maniera prevalente alla revisione dei pari anche per i settori delle scienze dure; la VQR italiana, recentemente conclusasi, ha operato nelle aree delle SSH per lo più affidandosi alla *peer review* o alla *informed peer review* (revisione dei pari integrata da dati relativi alle classifiche di riviste). Non a caso, entrambi gli esercizi di valutazione sono ritenuti estremamente

² Cfr. ad es. QUALITY INDICATORS (2011), p. 12 e p. 43; ARCHAMBAULT, VIGNOLA GAGNÉ (2004), p. 9 ss.: *Bibliometric methods are very useful for measuring the dissemination of knowledge in the natural sciences, but they are less effective in some applied fields, such as engineering. Their use in SSH evaluation poses specific problems; some solutions are available, but they need to be applied with considerable care. According to the research, this is due to structural differences in knowledge production and dissemination in the various fields. A number of scholars have highlighted fundamental differences between the scientific communication practices of scholars in the NSE and those in the SSH. These structural differences explain why the usefulness of bibliometric methods varies with the SSH discipline.*

dispendiosi.³ La cosa può anche essere ritenuta accettabile dal momento che tali esercizi si svolgono con una cadenza non troppo frequente.

La questione appare in una luce diversa quando si tratta di valutazioni d'ateneo o di struttura. In questo caso, il ricorso alla revisione dei pari diviene improponibile; è proprio per questo motivo che ora molte strutture cercano di adattare gli esiti della VQR piegandoli a fini non previsti, quali la valutazione interna di ateneo ai fini del reclutamento. In fondo, si tratta di risparmiare tempo e danaro.⁴

Inoltre, la *peer review* non gode di buona stampa: molti la considerano «soggettiva», in opposizione a una presunta «oggettività» dell'analisi citazionale, un argomento fallace ma diffuso.⁵ D'altro canto gli esiti della VQR sembrano dimostrare una eccessiva variabilità nei risultati della revisione dei pari, troppo spesso influenzata da appartenenze di scuola: ad esempio, nell'area 14 si è osservata una «notevole variabilità nei giudizi espressi sugli stessi prodotti, variabilità che ha sovraccaricato il lavoro dei gruppi di consenso».⁶ Nell'area 11 si è osservato come la valutazione dei pari si sia mostrata «con poche eccezioni facilmente spiegabili, meno generosa della seconda [bibliometrica]⁷ per quanto riguarda i prodotti eccellenti, con percentuali che in media si aggirano intorno al 15 % a fronte di una media di circa il 35% nella bibliometria, e allo stesso tempo meno propensa ad assegnare i prodotti valutati alla classe inferiore (Limitati), con percentuali medie al di sotto del 19% contro il 27% della bibliometria».⁸ La stessa congruenza fra valutazione del contenitore (rivista) e del contenuto (articolo) sembra in molti casi messa in discussione.⁹

³ Cfr. BACCINI, COIN, SIRILLI (2013), p. 55 ss.

⁴ Ciò non toglie che l'utilizzo di dati raccolti per la valutazione di strutture, nel corso di un esercizio di valutazione nazionale, per altri scopi (come le scelte dipartimentali in materia di reclutamento) è cosa che dovrebbe essere oggetto di attenta riflessione critica.

⁵ Sul punto cfr. GILLIES (2008), *passim*.

⁶ Rapporto finale di area. Gruppo di Esperti della Valutazione dell'Area 14 (GEV14), p. 64. http://www.anvur.org/rapporto/files/Area14/VQR2004-2010_Area14_RapportoFinale.pdf

⁷ Nell'area 11 esiste infatti una sottoarea bibliometrica (psicologia).

⁸ Rapporto finale di area. Gruppo di Esperti della Valutazione dell'Area 11 (GEV 11), p. 25. http://www.anvur.org/rapporto/files/Area11/VQR2004-2010_Area11_RapportoFinale.pdf

⁹ Si veda ad es. Rapporto finale di area. Gruppo di Esperti della Valutazione dell'Area 14 (GEV14), p. 32: «si può dire che la probabilità che un articolo pubblicato su una rivista di classe A sia valutato eccellente è leggermente più alta di quella di un articolo pubblicato su riviste classificate in fascia B

Ciò premesso, è chiaro che l'analisi bibliometrica sembra offrire una soluzione facile ed economica: il computo di indicatori di impatto (come ad esempio le citazioni) può essere automatizzato e realizzato in modo relativamente spedito. Di qui l'idea di estendere anche alle SSH strumenti di valutazione basati sulle citazioni.

In realtà, per la maggioranza delle SSH, l'analisi citazionale è ancora allo stadio embrionale, se non del tutto assente. In alcuni paesi essa è scartata a priori, per una serie di ragioni che cercherò di ricordare più sotto. Ciò significa che il progetto che ANVUR intende presentare ha una natura quasi pionieristica:¹⁰ il che dovrebbe di per sé indurre a cautela. La documentazione resa finora disponibile sembra invece essere caratterizzata più dall'entusiasmo proprio dell'esploratore di terre incognite che dalla consapevolezza dei rischi che ai quali ci si espone. Il che è ragione di allarme, tanto più che il progetto che sarà presentato nell'incontro del 20 gennaio sembrerebbe mirare ad avere immediate ricadute nelle procedure di valutazione (individuali e di struttura). Infatti nel documento relativo al progetto si specifica che i prodotti della ricerca indicizzati saranno assegnati a *subject categories* mantenendo «una classificazione, ad uso prevalentemente concorsuale e amministrativo, basata sulla normativa in essere ai fini della Abilitazione scientifica nazionale».¹¹ Ancora, relativamente all'analisi dei dati, si precisa che essa «attiene alle applicazioni necessarie per servire i singoli processi di valutazione, che possono essere fatti a diversi livelli (VQR, ASN, valutazioni interne degli atenei, ecc.)».¹² Infine, in relazione a fornitura e utilizzazione dei dati, è scritto che i destinatari potranno «utilizzare i dati e gli indicatori al fine di

o C o che la concordanza tra le valutazioni peer e la nostra classificazione è risultata piuttosto bassa, ma non trascurabile». http://www.anvur.org/rapporto/files/Area14/VQR2004-2010_Area14_RapportoFinale.pdf

¹⁰ Per un elenco di database che offrono una certa copertura delle SSH, ma che comunque non sono stati edificati con le finalità di valutazione che si propone ANVUR, cfr. ARCHAMBAULT, VIGNOLA GAGNÉ (2004), p. 5 ss. Si veda anche HICKS (2013), p. 76 ss.

¹¹ Specifiche preliminari per una base dati bibliometrica italiana nelle aree umanistiche e sociali, p. 10.

http://www.anvur.org/attachments/article/570/Allegato%201_Elementi%20architettura%20database.pdf

¹² Specifiche preliminari per una base dati bibliometrica italiana nelle aree umanistiche e sociali, p. 11.

svolgere valutazioni nei termini assegnati dalla legge (nel caso della Agenzia) o con procedure interne regolate dai rispettivi statuti, ad esempio per la allocazione di risorse di personale o finanziarie».¹³

Le procedure di valutazione individuali e di struttura, in particolare l'Abilitazione Nazionale (ASN) e la valutazione periodica della qualità della ricerca (VQR),¹⁴ ma così pure i futuri procedimenti di selezione concorsuale in sede locale, sono procedure amministrative, rette da regole proprie e sottoposte eventualmente, nel loro concreto svolgimento, al vaglio del giudice amministrativo. A giudizio di chi scrive, proporre di adottare indicatori bibliometrici di natura citazionale per le SSH nel contesto di tali procedure è segno di una debole conoscenza delle complessità di funzionamento della macchina amministrativa e della dimensione dei problemi posti dall'analisi bibliometrica delle SSH.

Sabino Cassese aveva lucidamente argomentato, tempo fa,¹⁵ circa il destino al quale ci avrebbero condotti procedure di analisi scientometrica di per sé fragili trasposte nel contesto di procedure amministrative. È francamente singolare che si ignori quel monito, insistendo sulla medesima linea, proprio ora che cominciano a vedersi gli esiti della prima tornata dell'ASN e che iniziano a delinarsi le dimensioni del contenzioso generato dall'adozione di criteri e parametri instabili e di difficile accertamento. Ma andiamo con ordine.

1. Architetture discutibili.

Ho già avuto modo di affermare più volte come – a mio avviso – l'architettura del sistema italiano della valutazione della ricerca e della formazione superiore non presenti i caratteri di solidità e buon disegno che pure sarebbero auspicabili.¹⁶ L'Agenzia è di fatto un'agenzia governativa e la previsione meramente procedurale e formale della nomina del Direttivo

¹³ Specifiche preliminari per una base dati bibliometrica italiana nelle aree umanistiche e sociali, p. 14.

¹⁴ L'esercizio nazionale di valutazione sulla base del quale viene allocata una quota rilevante del fondo di finanziamento ordinario alle università.

¹⁵ CASSESE (2012).

¹⁶ Cfr. ad es. BANFI (2012); BANFI (2013). BACCINI (2013).

con decreto del Presidente della Repubblica,¹⁷ spesso invocata a dimostrazione della terzietà e dell'indipendenza dell'Agenzia stessa¹⁸ può essere utilizzata come argomento risolutivo e di chiusura solo per chi ignora quali siano gli effettivi processi di scelta e quale sia la collocazione ordinamentale dell'Agenzia. Ma non è comunque questo neppure il principale dei problemi che affliggono ANVUR. L'Agenzia, infatti, cumula troppi compiti e attribuzioni, a fronte delle esigue competenze di cui può disporre il suo Direttivo. Per quanto possa valere come riferimento del momento, non si può non notare che la quasi totalità dei membri di quest'ultimo sono privi di esperienze pregresse circa l'esercizio di attività di valutazione su larga scala, per non parlare di attività di ricerca in materia scientometrica: ambiti specialistici che non possono, d'altro canto, qualificare la selezione di tutti i membri in qualsiasi momento.

Ma oltre i dati che attengono alle competenze disponibili, vi è un dato di sistema: la scelta di fare di ANVUR la sede che cumula compiti di valutazione della ricerca, della didattica, di accreditamento dei corsi e delle sedi; spetta ad ANVUR di intervenire anche sul mantenimento o la chiusura di intere sedi, finendo per divenire – nei fatti – un luogo nel quale non si estraggono dati utili al *policy-maker*, ma un luogo nel quale si definiscono le politiche stesse.¹⁹ Il cumulo di competenze attribuite ad ANVUR dalla l. 240/2010 e dai provvedimenti attuativi correlati ha finito per svuotare di competenze il Ministero (che pure dovrebbe vigilare sull'operato dell'Agenzia), trasferendole all'Agenzia stessa ma in nome di una presunta superiorità e affidabilità della tecnica rispetto alla politica. L'esito è una generalizzata deresponsabilizzazione, che rende difficile identificare i centri decisionali e attribuire la responsabilità di scelte politiche in materia di formazione terziaria e di ricerca.

In questo quadro è del tutto inopportuno che l'Agenzia decida di amplificare le proprie competenze, ad esempio proponendosi come luogo dove

¹⁷ DPR 76/2010, art. 8 c. 3: I componenti del Consiglio direttivo sono nominati con decreto del Presidente della Repubblica, su proposta del Ministro, sentite le competenti Commissioni parlamentari. Cfr. PINELLI (2011), p. 567 ss.

¹⁸ Cfr. ad es. BONACCORSI (2013).

¹⁹ In questo quadro, non manca neppure il salotto scientifico-letterario dell'Agenzia: «I mercoledì dell'ANVUR». http://www.anvur.org/index.php?option=com_content&view=article&id=515:i-mercoled%C3%AC-dell-anvur-it&catid=48&Itemid=363&lang=it

sviluppare ricerche ed esperimenti scientometrici: la scientometria è una disciplina delicata, per le conseguenze che essa può comportare nel ridisegnare il sistema di ricerca di un intero paese. Proprio per questa ragione, essa deve essere lasciata alla comunità scientifica e non fatta in casa da un ristretto consesso di persone interne a un'agenzia ministeriale. Analogamente, non pare opportuno che l'Agenzia, un'entità allo stesso tempo troppo forte (per l'insieme delle attribuzioni e per l'impatto sulle *policies*) e troppo debole (per esiguità di risorse e di competenze), si faccia promotrice dell'edificazione *in-house* di un database bibliometrico-citazionale, sia pur appoggiandosi a risorse esterne per quanto riguarda la sua gestione.²⁰

La bibliometria come funzione, allocabile presso le amministrazioni dello Stato è un fenomeno sostanzialmente ignoto al panorama della valutazione internazionale: può non piacere che l'analisi dell'impatto citazionale sia affidata di norma a due dei maggiori *players* editoriali al mondo (Elsevier e Thomson-Reuters); essi tuttavia dispongono del *know-how* necessario allo sviluppo delle proprie attività e soprattutto non costituiscono parte integrante del medesimo organismo incaricato di produrre gli esiti della valutazione. Un sistema sano e equilibrato di valutazione prevede che i compiti siano distribuiti, proprio per evitare che si sovrappongano compiti che devono essere tenuti separati: esistono perciò fornitori di dati bibliometrici (ISI, Scopus/WoS ecc.), elaboratori ed interpreti di dati (agenzie di valutazione) che dovrebbero operare sulla base di regole non definite da loro stessi, ma costruite all'interno di un ampio dibattito che includa le comunità scientifiche e accademiche e infine soggetti che sulla base di quei dati organizzati e interpretati, prendono le proprie decisioni (i *policy makers*). Tale segmentazione evita che un singolo organismo diventi il fulcro (scientifico, tecnologico, valutativo e politico) del sistema della formazione e della ricerca di un intero paese.

In particolare, appare poco o per nulla opportuno che ANVUR sia il progettista e gestore ultimo di un database bibliometrico per le SSH. In un Paese maturo, per quel che riguarda la cultura della valutazione, sul punto

²⁰ Specifiche preliminari per una base dati bibliometrica italiana nelle aree umanistiche e sociali, p. 14.

non vi sarebbe pressoché discussione. Essendo un paese immaturo – e il livello di buona parte della bibliometria e della scientometria promossa da ANVUR lo dimostra²¹ – si mette in agenda un ulteriore accentramento di competenze intorno all’Agenzia. Eppure, ci sarebbero buoni motivi, oltre a quelli esposti finora, per allarmarsi seriamente. Segnali, che indicano come la sovrapposizione dei compiti sia destinata a produrre gli effetti perversi che ci si possono attendere da un errato disegno istituzionale.

Basti considerare i seguenti aspetti: un progetto «pilota» viene presentato dall’Agenzia presso il CNR. Si auspica la condivisione e il dibattito da parte della comunità scientifica. Del resto, anche D. Hicks ricordava che per quanto concerne la valutazione bibliometrica delle SSH *any successful infrastructure will need to productively engage with the scholarly community*.²² Eppure l’architettura, la gestione, perfino il *business model* dell’operazione sono già sostanzialmente definiti, così come ne è definito lo scopo. Cosa resta da discutere? E soprattutto con chi si intende discutere? I primi interlocutori identificati non sembrano rappresentativi della comunità scientifica, ma solo di alcuni suoi attori direttamente interessati e forse in positivo.²³

Volendo si sarebbero potuti utilizzare altri strumenti, prendendo ad esempio la consultazione pubblica avviata tempo fa dal CUN sulla costituzione di ANPREPS.

Soprattutto, ci si deve augurare che ogni scelta in materia sia cauta, ponderata e «lasciata» discutere dalle e nelle comunità, al di là di quelle logiche emergenziali che hanno sin qui caratterizzato la valutazione. L’introduzione delle mediane per la selezione di commissari e candidati è l’espressione di ciò a cui può condurre la valutazione nell’emergenza e dell’emergenza: frutto della convinzione (indimostrata) che la produzione scientifica di una larga parte dei ricercatori italiani fosse nulla o gravemente carente, sicché il 50% di costoro poteva essere escluso dalle procedure.²⁴

²¹ Per una rapida rassegna si veda BANFI, DE NICOLAO (2013).

²² HICKS, WANG (2009), p. 3.

²³ L’incontro fissato per il 20 gennaio 2014 è a inviti e aperto esclusivamente ai presidenti delle società scientifiche delle SSH, e ai direttori e agli editori di riviste di fascia A.

²⁴ Come è noto l’esclusione è stata automatica per quanto concerne la selezione dei commissari; diverso il caso dei candidati.

Analogamente, sulla base del presupposto che le SSH italiane siano provinciali, periferiche, di scarso impatto a livello internazionale (presupposto non dimostrato e nemmeno facile da dimostrare), si vuole incrementare la produttività dei ricercatori e potenziare il ruolo delle riviste di fascia A attraverso strumenti bibliometrici. Tali riviste – almeno in alcuni settori – rappresenterebbero infatti la punta di diamante quanto a internazionalizzazione e qualità della ricerca pubblicata. Non è in alcun modo possibile verificare se ciò sia sempre vero, mentre al contrario sono ben noti i problemi generati dai *rankings* di riviste.²⁵ Tanto meno è opportuno mescolare analisi citazionale (del tutto quantitativa) e attribuzione di *ranks* (operazione compiuta attraverso la *peer review* dei cosiddetti gruppi di esperti della valutazione – GEV). Di tutto questo sarebbe opportuno discutere a fondo, ma – come si diceva – le logiche emergenziali portano a considerare un'autentica discussione come una perdita di tempo, se non – peggio – come lo strumento grazie al quale le parti meno qualificate dell'accademia intendono frenare la valutazione. Non è il caso di ripetere errori già compiuti.²⁶

2. Bibliometria e SSH: una questione aperta.

Non è questa la sede dove ricordare il complesso di problemi tecnici e teorici posti dall'analisi citazionale nelle SSH né tutte le ragioni (largamente riconosciute a livello scientifico ma anche da organizzazioni e riviste internazionali)²⁷ che depongono contro un uso indiscriminato della bibliometria per la valutazione individuale. Mi limito a sollevare due punti: in primo luogo, la sistematica sottovalutazione della difficoltà del compito che ANVUR si propone, a maggior ragione quando ci si propone di far ricadere gli esiti di tale compito quali elementi di valutazione all'interno di procedure amministrative. In secondo luogo, il disinteresse dimostrato dagli estensori del documento preliminare per la questione cruciale della rilevanza statistica delle citazioni nei settori delle SSH.

²⁵ I lettori mi perdoneranno se mi permetto, a questo proposito, di rinviare al mio BANFI (2013.2).

²⁶ A proposito di opposizione alla valutazione quantitativa nelle SSH e di «costruzione del consenso» cfr. HUG, OCHSNER, DANIEL (2012).

²⁷ Mi riferisco in particolare alla *San Francisco Declaration On Research Assessment* (DORA). <http://am.ascb.org/dora/>

La *roadmap* prevista da ANVUR prevede che il costituendo database includa in primo luogo – in una fase «sperimentale» – le sole riviste censite in fascia A ai fini dell’ASN e della VQR. Alla conclusione della fase sperimentale si procederà alla definizione di indicatori bibliometrici.²⁸ Si pone dunque la questione degli stili citazionali delle SSH che notoriamente sono molto diversi rispetto a quelli delle scienze dure e spesso disuniformi fra di loro. In particolare, solo raramente gli umanisti fanno ricorso a bibliografie collocate alla fine dei loro articoli. Ci si propone dunque di costruire un software che sia in grado – anche attraverso l’autoapprendimento – di «acquisire un numero sufficientemente elevato di stili citazionali».

Ora, viene spontaneo chiedersi che significato possa avere una base dati già costruita su di un numero estremamente limitato di fonti (le monografie sono escluse, e così pure tutte le riviste censite come scientifiche ma non di fascia A) e che per di più sarà in grado di identificare un numero «sufficientemente elevato» di stili citazionali all’interno del proprio limitato novero di fonti. Quanto dovrà essere elevato tale numero per essere sufficiente? E che uso si potrà mai fare dei dati estratti, dal momento che essi rappresenteranno solo una frazione del novero complessivo di citazioni nelle SSH? Nel documento di ANVUR relativo al progetto si fa cenno alla possibilità di interfacciarsi con altre fonti per migliorare l’accuratezza nell’estrazione dei dati: per esempio attraverso ORCID.²⁹ Tuttavia, viene spontaneo chiedersi quanti siano gli umanisti e scienziati sociali dotati di un profilo ORCID. Dati di esperienza possono essere fuorvianti, ma – per fare un solo esempio - chi scrive è pressoché certo di essere l’unico nel proprio settore disciplinare.

ANVUR si propone un livello di accuratezza della base dati prossimo al 100%: eppure, come giustamente ricorda l’estensore del documento programmatico, neppure WoS e Scopus, che pure operano da decenni con risorse colossali e su di un materiale che per lo più pone molti meno

²⁸ Specifiche preliminari per una base dati bibliometrica italiana nelle aree umanistiche e sociali, p. 4.

²⁹ <http://orcid.org/> «ORCID provides a persistent digital identifier that distinguishes you from every other researcher and, through integration in key research workflows such as manuscript and grant submission, supports automated linkages between you and your professional activities ensuring that your work is recognized».

problemi di estrazione dei dati restituiscono risultati privi di errori: le conseguenze di tali errori si sono viste recentemente anche in occasione dell'ASN e del computo delle mediane. Occorreranno allora interventi manuali di correzione: il che pone un'ulteriore questione.

Se la scelta dell'analisi bibliometrica dovrebbe giovare a ridurre i costi dell'attività di valutazione, quali saranno i costi complessivi, diretti e indiretti, dell'operazione che si va prospettando? Le SSH in larghissima parte non necessitano dell'allocazione di finanziamenti per la ricerca di dimensioni particolarmente rilevanti: salvo rare eccezioni non occorrono laboratori, macchinari e così via. Sarà dunque interessante verificare il rapporto fra il costo complessivo dell'operazione avviata da ANVUR e una stima dei finanziamenti complessivamente assegnati alle SSH. Della questione dei costi l'Agenzia pare essere ben consapevole: infatti, con il fine di contenerli, auspica che sia la collaborazione volontaria di autori ed editori a garantire la qualità del database.

Tornando ai problemi pratici posti dall'analisi citazionale nelle scienze umane e sociali, ANVUR riconosce che le citazioni spesso non possono essere trattate come una semplice misura dell'impatto – a differenza di quanto accade, per lo più, nelle scienze dure. Infatti sono elencate citazioni per «adesione», per «tributo», per «rassegna», per «dissenso», per «rituale accademico».³⁰ Sarà interessante capire se e come l'Agenzia distinguerà le citazioni secondo categorie interpretative. Un'operazione che non ha alcuna possibilità di successo neanche se eseguita manualmente.

La verità è che – salvo specifiche eccezioni – la macroarea delle SSH non è pronta per un'analisi citazionale che non sia **estremamente grossolana**. Perché gli stili citazionali non sono uniformi e perché la tecnica delle citazioni è ispirata a un'antica tradizione che rende malagevole identificare univocamente il dato citazionale con un dato di impatto.³¹ Indubbiamente, si tratta di problemi non irrisolvibili: occorrerebbe però che gli editori e le comunità scientifiche si adattassero a un nuovo modo di procedere nel citare. Col tempo ciò è certamente fattibile. Molto tempo, però: non si cambiano

³⁰ Specifiche preliminari per una base dati bibliometrica italiana nelle aree umanistiche e sociali, p. 13.

³¹ Sul punto cfr. GRAFTON (1994).

tradizioni secolari dall'oggi al domani attraverso iniziative verticistiche di sapore vagamente sovietico. Ma ammettiamo pure che tutti questi problemi di natura tecnica siano superabili. Ne resta un altro che, ad avviso di chi scrive, è insuperabile.

Infatti, nelle SSH esistono settori che sono per più versi accostabili ai settori delle scienze dure: penso ad esempio alla linguistica e alla glottologia, alla filosofia della scienza e del linguaggio, oltre che – ovviamente – alla psicologia. Generalizzando, si potrebbe dire che tali settori si caratterizzano per forte internazionalizzazione, largo uso dell'inglese, maggior produzione di articoli e minor produzione di studi monografici. Per questo motivo, queste discipline si prestano meglio all'analisi citazionale: non solo per una questione di «stile» della ricerca, ma perché esse si muovono all'interno di un'ampia comunità scientifica internazionale. D'altra parte esistono settori che per ovvie ragioni, che nulla hanno a che fare con la qualità della ricerca prodotta, hanno un'audience quasi esclusivamente nazionale (se ne trovano svariati fra le scienze giuridiche, ma non solo) o che hanno una dimensione estremamente ridotta, quanto a numero di ricercatori, sia a livello nazionale che sovranazionale. Per tutte queste discipline, che costituiscono una parte tutt'altro che irrilevante delle SSH, una macroarea tutt'altro che omogenea,³² il numero di citazioni potenziali (in uscita e in entrata) è per forza di cose scarso. Che senso ha, dunque, computare indicatori citazionali quando essi si fonderanno su di un numero così basso di citazioni complessive da essere del tutto privi di rilevanza statistica? Si usa dire che le distorsioni proprie dei dati citazionali sono rese meno dirompenti dai «grandi numeri». Ma vi sono davvero «grandi numeri» nelle SSH? La questione è tanto più rilevante, in quanto proprio a causa delle finalità attribuite alla banca dati (ASN e VQR), i prodotti da cui saranno estratte le citazioni saranno catalogati in primo luogo secondo l'attuale configurazione per settori disciplinari e concorsuali.³³ Avremo così settori disciplinari per i quali il numero di citazioni complessive è irrilevante e settori concorsuali all'interno dei quali si confronteranno settori con un maggiore o minor numero di citazioni

³² GIMÉNEZ-TOLEDO, ROMÁN-ROMÁN (2009).

³³ Specifiche preliminari per una base dati bibliometrica italiana nelle aree umanistiche e sociali, p. 9 ss.

complessive, a tutto danno dei secondi. Non poteva mancare, in questo contesto, un cenno da parte dell'estensore del documento ANVUR alla sciagurata pratica della normalizzazione per settore disciplinare. Vale la pena di ribadirlo ancora una volta: se i numeri sono troppo piccoli, non c'è normalizzazione che tenga. I dati sono semplicemente inutili.

Il quadro della situazione è dunque il seguente: ci si propone di creare un database con scarsa copertura, con gravissimi problemi di accuratezza, risolvibili – forse – solo a prezzo di un ingente sforzo economico, che comunque in molti casi restituirà informazioni irrilevanti e dunque del tutto inutilizzabili sia ai fini della valutazione sia individuale che – con ogni probabilità – di struttura.

3. Un incentivo a comportamenti non etici?

La recente irruzione di metriche quantitative nel reclutamento accademico e nella valutazione di strutture ha già prodotto i suoi effetti nefasti sull'etica dei ricercatori. In un quadro nel quale nessuno si è voluto preoccupare della diffusione di comportamenti opportunistici e abusivi e dell'urgente necessità di porvi rimedio, si sono moltiplicati fenomeni come il *salami slicing* (la moltiplicazione dei prodotti di ricerca attraverso la loro frammentazione: una scorsa ai *curricula* per l'ASN di commissari e candidati delle SSH è sufficiente a dimostrarlo) e gli schemi citazionali miranti a gonfiare, nelle scienze dure, i punteggi di riviste e dei singoli autori che su di esse pubblicano.³⁴

In questo quadro, l'Agenzia promuove un database bibliometrico per le SSH fondato in primo luogo sulle riviste di fascia A. E' evidente che, nelle intenzioni dell'ANVUR, si intende misurare in primo luogo le eccellenze. La scelta di ANVUR, però, rischia di tradursi in un incremento del potere del quale dispongono le direzioni di tali riviste, che sono state selezionate come «eccellenti» sulla base della *peer review* e non certo dei dati citazionali. Viene così a crearsi un cortocircuito fra quantitativo e qualitativo, dal quale non c'è da aspettarsi nulla di buono sotto il profilo dell'etica della ricerca e delle pubblicazioni. Tanto più che l'Agenzia si propone di «pesare» diversamente le citazioni sulla base della loro provenienza: da riviste di fascia A o da riviste «internazionali» (qualsiasi

³⁴ Cfr. ad es. NATURE 500 (29 Aug. 2013), p. 511.

cosa si intenda con questo termine ancora in attesa di definizione). Dobbiamo dunque supporre che ci troveremo nella situazione in cui una citazione proveniente da una rivista definita «eccellente» sulla base della revisione dei pari peserà di più di una citazione proveniente da una rivista solo «scientifica».³⁵ Un marchingegno mai visto prima,³⁶ che può spiegarsi solo con la pervicace volontà dell’Agenzia di promuovere gruppi e settori che lo stesso direttivo considera «eccellenti», unita a una perdurante incompetenza bibliometrica e scientometrica.³⁷ Vi sarebbe da chiedersi se questo sia il fine della valutazione: quello di conformare il sapere.

In queste scelte eminentemente dirigistiche, grazie alle quali l’Agenzia si propone di ridisegnare il panorama delle SSH italiane, si nasconde ancora una volta, sotto la maschera tecnocratica, una scelta di *policy*, che esautora le comunità scientifiche e ne mortifica l’autonomia. Ad esempio, il database non prevede l’indicizzazione delle monografie: il che si tradurrà, verosimilmente, nell’abbandono da parte di molti ricercatori, specie i più giovani, di quel genere letterario, benché esso sia il genere letterario per eccellenza di buona parte delle SSH. Si può anche pensare che sia meglio produrre articoli, magari in lingua inglese, piuttosto che monografie. Questa però è una scelta che spetta alle discipline e ai ricercatori, e non può essere l’effetto di scelte dell’Agenzia congegnate per riorientare il mondo della ricerca secondo gli indirizzi che più piacciono a questo o quel componente il direttivo ANVUR. Ribadisco: l’agenzia deve valutare secondo criteri definiti altrove, non «fare» le politiche della ricerca. Eppure, il dirigismo *à la* Lysenko dell’Agenzia è tale da indurre a una sistematica sottovalutazione dei rischi impliciti nell’operazione da essa lanciata anche per l’etica della ricerca: è evidente che – specie in settori poco numerosi – basteranno pochi

³⁵ Specifiche preliminari per una base dati bibliometrica italiana nelle aree umanistiche e sociali, p. 12 ss.

³⁶ Tentativi di pesatura delle citazioni per «prestigio» delle sedi di pubblicazione sono stati esaminati in via teorica, ma non sono mai stati pensati al di fuori dell’ambito citazionale: intendendosi con popolarità il numero delle citazioni di un autore e con prestigio il numero di citazioni provenienti da riviste molto citate. Cfr. DING, CRONIN (2010).

³⁷ Sulla costruzione di indicatori comprendenti misure di impatto e prestigio (cosa ben diversa da quanto proposto da ANVUR) si vedano WEST, RICH (2012). Gli autori ricordano comunque quanto segue: «we do not recommend using these criteria uncritically to generate a ranking of journals that “count” and “do not count” since all of these data points can be skewed, manipulated, or changed from year to year».

giorni per creare cartelli citazionali o per incardinare pratiche abusive al fine di massimizzare i risultati della propria produzione scientifica. Tanto più che, in assenza dell'ANPREPS, l'Agenzia non intende neppure distinguere fra i prodotti di ricerca pubblicati sulle riviste. Note, recensioni, lettere, articoli, peseranno lo stesso (del resto non è semplice operare una classificazione o distinzione): mezza paginetta di recensione su una rivista «eccellente» varrà presto come oro; e questa non vuole essere solo una metafora, visto quanto sta accadendo in Paesi che hanno adottato simili procedure.³⁸

Se ne gioverà la ricerca italiana? In questo quadro, l'estensore del documento di ANVUR dedica solo un paio di righe al possibile sviluppo «in futuro» di indicatori «per ambiti di ricerca più di dettaglio, per evitare che per ciascun settore concorsuale siano privilegiate le pubblicazioni che affrontano i temi di ricerca più popolari». Per il resto, silenzio: neppure un cenno al problema, notissimo, delle autocitazioni. Anzi, in una sorta di infantile entusiasmo per il «nuovo», che si deve supporre dovuto al senso di inferiorità diffuso fra certi umanisti nei confronti dei loro colleghi delle scienze dure, unito al fascino che indicatori quantitativi esercitano su chi non è avvezzo a maneggiare numeri, ci si propone di esplorare la via di Altmetrics,³⁹ ossia di considerare la creazione di indicatori fondati sul numero di letture di articoli *on-line*, sul numero di *downloads*, perfino sui *social networks* per ricercatori (come ad esempio *academia.edu* o *Researchgate*).

Se le cose stanno così, meglio Publish or Perish e Google Scholar. Per avere un risultato scadente, inutilizzabile, e che si spera resti inutilizzato, è inutile sprecare soldi pubblici.

4. Conclusioni e suggerimenti

Il bilancio costi-benefici dell'iniziativa di ANVUR appare profondamente negativo. I costi per ottenere una base dati comunque parziale ma almeno limitatamente attendibile saranno elevati. I benefici, sostanzialmente nulli. Gli svantaggi, sulla qualità della produzione scientifica e sull'etica delle comunità di ricerca, elevati. Così pure gli effetti indiretti di riorientamento

³⁸ ŠIPKA (2012).

³⁹ <http://altmetrics.org/manifesto/>

delle attività di ricerca verso filoni *mainstream*. Alla luce di queste considerazioni e delle informazioni finora rese disponibili da ANVUR il progetto del database bibliometrico appare inaccettabile. Le comunità scientifiche hanno molte e fondate le ragioni per avversarlo fosse solo per scongiurare i danni evidenziati in questo articolo. In particolare, vale la pena di ribadire ancora una volta come le procedure di valutazione non debbano trasformarsi in operazioni etico-pedagogiche: non spetta all’Agenzia orientare il comportamento delle comunità scientifiche secondo i dettami di proprie tavole della legge: una bibliometria di Stato che si prefigga di promuovere la «buona scienza» separandola da quella «cattiva» rappresenta un pericolo per la scienza italiana che ricorda in modo sinistro le politiche della ricerca promosse a suo tempo da Lysenko. Tanto più che il database concepito dall’Agenzia non solo per lungo tempo non sarà in grado di offrire che una scarsissima copertura della ricerca delle SSH, ma probabilmente non sarà mai completo. Viste le finalità pedagogiche sottese all’operazione (in buona sostanza l’ulteriore promozione delle riviste di classe A come luogo eccellente di pubblicazione) non mi stupirei nemmeno se l’ampiamiento agli altri periodici, per ora annunciato per il futuro, non avvenisse mai. Ma in ogni caso, che fine faranno le citazioni confinate all’interno di riviste non censite? Saranno prive di valore e considerate in via automatica appartenenti alla «cattiva scienza»? Eppure è tutt’altro che improbabile che anche al di fuori delle riviste di fascia A si troveranno studi meritevoli ed eccellenti.

Tutto ciò significa che non sia possibile procedere a una valutazione delle SSH attraverso indicatori almeno parzialmente quantitativi? Certamente no. Occorre però che si definisca un percorso chiaro e condiviso, orientato, come ha ben argomentato, fra gli altri, Paola Galimberti, verso la definizione di indicatori multipli (non solo citazionali) in grado di abbracciare la complessità delle SSH e anche la terza missione spesso svolta da chi opera in tali settori.⁴⁰

A questo scopo occorre un progetto pilota ben costruito, approfonditamente discusso, che tenga conto della letteratura scientometrica internazionale in materia, e che sia sottoposto a test per un periodo di tempo sufficientemente lungo da poter essere valutato e nel caso corretto. Ci vorrebbero anni,

⁴⁰ GALIMBERTI (2013); OCHSNER, HUG, DANIEL (2012).

probabilmente tre o quattro. In tal modo l'Italia potrebbe davvero uscire dalla sua posizione di «fanalino di coda» della valutazione internazionale. Sarebbe una scelta responsabile e di buon senso: purtroppo essa cozza contro l'urgenza dei «rivoluzionari dall'alto» ai quali una cattiva regolazione ha consegnato le redini del sistema della valutazione dell'università e della ricerca.

RIFERIMENTI BIBLIOGRAFICI

ARCHAMBAULT, VIGNOLA GAGNÉ (2004): E. Archambault, E. Vignola Gagné, The Use of Bibliometrics in the Social Sciences and Humanities, Science Metrix Final Report 2004. (http://www.science-metrix.com/pdf/SM_2004_008_SSHRC_Bibliometrics_Social_Science.pdf)

BACCINI (2013): A. Baccini, Come e perché ridisegnare la valutazione, Il Mulino 1.2013, pp. 80-87.

BACCINI, COIN, SIRILLI (2013): A. Baccini, F. Coin, G. Sirilli, Costi e benefici della valutazione della ricerca e della didattica, Paradoxa 7.2, pp. 49-61.

BANFI (2012): A. Banfi, Salvare la valutazione dall’Agenzia di Valutazione?, Federalismi.it, Rivista di diritto pubblico italiano, comunitario e comparato, 22 (2012).

BANFI (2013): A. Banfi, Prima che la nave affondi: un rapido bilancio della riforma dell’Università e qualche possibile intervento correttivo, Giornale di Diritto Amministrativo 5/2013, pp. 548-559.

BANFI (2013.2): A. Banfi, Il resistibile fascino delle classifiche di riviste, RIV – Rassegna Italiana di Valutazione 51 (2011).

BANFI, DE NICOLAO (2013): A. Banfi, G. De Nicolao, Potenzialità e limiti degli indici bibliometrici nella valutazione della ricerca scientifica, Paradoxa 7.2, pp. 34-48.

BONACCORSI (2013): A. Bonaccorsi, La valutazione di Bertoldo, Il Mulino 2.2013, pp. 257-265.

CASSESE (2012): S. Cassese, L’Anvur ha ucciso la valutazione, viva la valutazione, Il Mulino 1.2013, pp. 73-79.

DING, CRONIN (2010): Y. Ding, B. Cronin, Popular and/or Prestigious? Measures of Scholarly Esteem. <http://arxiv.org/abs/1012.4871>

GALIMBERTI (2013): P. Galimberti, Valutazione e scienze umane: limiti delle attuali metodologie e prospettive future, Astrid Rassegna 191 (20/2013).

GILLIES (2008): D. Gillies, How Should Research be Organized?, London 2008.

GIMÉNEZ-TOLEDO, ROMÁN-ROMÁN (2009): E. Giménez-Toledo, A. Román-Román, Assessment of humanities and social sciences monographs

through their publishers: a review and a study towards a model of evaluation, *Research Evaluation* 18, 3.

GRAFTON (1994): A. Grafton, The Footnote from De Thou to Ranke, *History and Theory* 33.4, pp. 53-76.

HICKS, WANG (2009): D. Hicks, J. Wang, Towards a Bibliometric Database for the Social Sciences and Humanities, From the Selected Works of D. Hicks.

http://works.bepress.com/cgi/viewcontent.cgi?article=1017&context=diana_hicks

HICKS (2013): D. Hicks, One size doesn't fit all: On the co-evolution of national evaluation systems and social science publishing, *Confero* 1.1, pp. 67-90.

HUG, OCHSNER, DANIEL (2012): S.E. Hug, M. Ochsner, H.D. Daniel, A Framework to Explore and Develop Criteria for Assessing Research Quality in the Humanities, *International Journal for Education Law and Policy* (in press).

OCHSNER, HUG, DANIEL (2012): M. Ochsner, S.E. Hug, H.D. Daniel, Indicators for Research Quality for Evaluation of Humanities Research: Opportunities and Limitations, *Bibliometrie – Praxis und Forschung* 1 (2012).

PINELLI (2011): C. Pinelli, Autonomia universitaria, libertà della scienza e valutazione dell'attività scientifica, *MUNUS* 3, pp. 567-588.

QUALITY INDICATORS (2011): The Royal Netherlands Academy of Arts and Sciences, Quality Indicators for Research in the Humanities, 2011. (<https://www.knaw.nl/shared/resources/actueel/publicaties/pdf/20111024.pdf>)

ŠIPKA (2012): P. ŠIPKA, Legitimacy of citations in predatory publishing: the case of proliferation of papers by Serbian authors in two Bosnian WoS-indexed journals. CEES Occasional Paper Series, 2012-12-2. (<http://www.ceon.rs/ops/12122>)

WEST, RICH (2012): R.E. West, P.J. Rich, Rigor, Impact and Prestige: A Proposed Framework for Evaluating Scholarly Publications, *Innovative Higher Education* 37.5.

