

Problemi e prospettive per la valutazione della ricerca in Italia

Paolo Rossi

Dipartimento di Fisica "E. Fermi" dell'Università di Pisa

Consiglio Universitario Nazionale (CUN)

La ricerca in Italia, per motivi riconducibili in larga misura alla struttura e all'articolazione del sistema produttivo, è purtroppo confinata quasi totalmente al settore pubblico, le cui risorse sono peraltro scarse e per di più decrescenti. Il posizionamento complessivamente buono della ricerca italiana (attestato da studi internazionali basati su metodologie oggettive e indipendenti)¹ non è tuttavia omogeneo tra le differenti aree disciplinari né tantomeno può considerarsi assicurato a tempo indeterminato se non verranno messe in atto adeguate politiche di sostegno. Tali politiche dovranno necessariamente comportare una stretta combinazione tra una maggior disponibilità di risorse (umane e finanziarie) e un aumento di efficienza nel loro utilizzo. Premessa indispensabile a tale aumento di efficienza è certamente l'acquisizione di una reale capacità di valutazione *ex post* dei risultati, sulla cui base impostare interventi premiali o, dove necessario, disincentivanti.

Il tema della valutazione ha quindi assunto negli ultimi tempi una centralità assoluta nel contesto delle problematiche relative al sistema universitario italiano. In effetti tutte le attività svolte nel contesto universitario, dalla didattica alla gestione amministrativa, e tutti i soggetti, individuali e collettivi, in esso operanti, sono potenzialmente oggetto di scrutinio a fini valutativi, ma in questa sede si intende focalizzare l'attenzione sui problemi specificamente collegati alla valutazione della ricerca.

L'aspirazione a metodologie valutative che possano a buon diritto definirsi "obiettive" e molto forte e non soltanto in Italia, ma la validità delle metodologie quantitative finora proposte (di natura prevalentemente bibliometrica) è tuttavia ancora oggetto di un vasto dibattito. Sarà quindi in questa fase necessario formulare ipotesi di lavoro, al momento evidentemente provvisorie, e metterle concretamente alla prova per misurarne la validità, sempre mantenendo una chiara consapevolezza del fatto che nel breve-medio periodo, a fronte di evidenze negative, occorrerà operare revisioni anche sostanziali della impostazione inizialmente adottata. Ogni norma che abbia carattere prescrittivo dovrebbe essere considerata come provvisoria, e i suoi effetti dovrebbero essere attentamente monitorati per cogliere gli eventuali segnali di reazioni perniciose da parte del sistema della ricerca, quali lo sviluppo di comportamenti opportunistici o l'abbandono di linee di ricerca culturalmente e scientificamente importanti ma poco remunerative in termini di valutazione "obiettiva".

Sembra utile in primo luogo cercare di identificare le principali opzioni metodologiche e i loro possibili domini di applicazione, notando che in contesti differenti possono risultare adeguate opzioni anche molto diverse tra loro. La varietà delle situazioni cui può riferirsi la valutazione della ricerca può essere sommariamente classificata secondo i seguenti principali parametri:^{2 3}

- a) finalità della valutazione (finanziamenti, retribuzioni, reclutamento, carriera accademica)
- b) soggetti della valutazione (individui, dipartimenti, atenei, enti)
- c) metodologia valutativa (*peer review*, criteri bibliometrici, metodi misti)
- d) oggetti della valutazione (articoli, monografie, brevetti, altri prodotti di ricerca)
- e) grandi aree disciplinari (scientifico-tecnologica, scienze della vita, umanistico-sociale)

Finalità della valutazione

In merito alle finalità della valutazione una prima e fondamentale distinzione è volta a separare le azioni valutative finalizzate alla pura e semplice verifica di superamento di prefissati requisiti minimi di produttività scientifica, individuale o collettiva, dalle azioni che hanno invece come obiettivo quello di formulare graduatorie (*ranking*), in genere mediante attribuzione di punteggi (*rating*), al fine di reclutamento, avanzamento di carriera, finanziamento di progetti, ripartizione di risorse e più in generale con obiettivi di incentivazione e premialità.

In linea di principio il primo aspetto, pur denso di importanti conseguenze pratiche, sembrerebbe potersi prestare abbastanza facilmente alla definizione di criteri di valutazione basati su parametri di natura prevalentemente quantitativa, in quanto il carattere non comparativo del giudizio potrebbe ridurre la valutazione alla verifica del superamento di una soglia quantitativa relativa a uno o più indicatori di produttività scientifica, ancorché certamente sia gli indicatori che la soglia debbano essere scelti con la massima attenzione all'accuratezza e all'efficacia.

In pratica la questione tende a complicarsi sotto diversi profili. In primo luogo, quando si voglia associare a questo tipo di valutazione qualche obiettivo strategico (ad esempio l'innalzamento della qualità media), sembra diventare inevitabile associare in qualche modo i valori di soglia a valori medi attualmente presenti nel sistema (o in qualcuna delle sue componenti). Tale associazione non è tuttavia scevra da rischi, in quanto la determinazione "storica" di tali valori medi e in certi casi la loro stessa definizione "algebrica" possono dipendere in modo talvolta cruciale dalle differenti dinamiche e modalità organizzative presenti nei vari settori disciplinari. Solo a titolo di esempio, si consideri l'importanza che può avere la presenza di una struttura fortemente gerarchica della ricerca, che porterà all'innalzamento dei valori medi dei gruppi dirigenti fino a livelli difficilmente raggiungibili per chi di quei gruppi dirigenti non faccia ancora parte, oppure la rilevanza di una corretta stima del grado di proprietà individuale del prodotto di una collaborazione in quei settori in cui la dimensione delle collaborazioni può variare da pochissime unità a molte centinaia, con gli ovvi effetti differenziati di produttività.

Non bisogna poi trascurare il fatto che la richiesta di superare valori medi attualmente esistenti può essere penalizzante nei settori in cui il livello è già coerente con gli *standard* internazionali (e quindi non richiede in realtà di essere sistematicamente superato, mentre sarebbe ragionevole attestare la soglia su valori inferiori, anche se non troppo, alla media), e viceversa nei settori in cui la qualità media attuale è troppo bassa un criterio legato alla media non permetterebbe comunque un rapido miglioramento della qualità.

Tutte queste considerazioni sembrano suggerire l'importanza di rifuggire dalle facili generalizzazioni, e anche qualora si decida di fare riferimento prevalente a criteri e parametri di natura quantitativa occorrerà calibrare le soglie caso per caso tenendo conto della specificità dei settori di ricerca, oltre che della maggiore o minore facilità di pervenire a una determinazione empirica attendibile dei valori di parametri che non in tutte le aree sono a tutt'oggi adeguatamente repertoriati e resi reperibili. Anche il riferimento a basi di dati bibliometrici internazionali non è sempre garanzia di completezza o di piena attendibilità anche solo quantitativa.

Se già la valutazione individuale non comparativa presenta, come abbiamo visto, numerosi aspetti problematici, la valutazione comparativa vera e propria richiede un'analisi ancor più sofisticata e impegnativa, che dovrà necessariamente fare attento e distinto riferimento a ognuno dei parametri generali inizialmente elencati.

Soggetti della valutazione e metodologia valutativa

Restando fin d'ora inteso che anche la natura degli oggetti della valutazione e l'area disciplinare di riferimento comportano distinzioni talvolta cruciali e richiedono analisi appropriate e proposte specifiche, sembra tuttavia opportuno organizzare il nucleo della discussione sulla base di due soli grandi assi di riferimento, quello relativo ai soggetti della valutazione, che possono variare dal singolo studioso ai grandi numeri di un ateneo, e quello relativo alla metodologia, che può variare con una certa continuità dalla valutazione cosiddetta "calda" e puramente qualitativa (*peer review*) a quella "fredda" e prevalentemente quantitativa degli indicatori bibliometrici.

Sembra evidente che intercorre in generale una profonda differenza tra la valutazione dei singoli, ai fini di incentivazione, di finanziamento della ricerca, di progressione stipendiale e di carriera, e la valutazione aggregata (unità operative, dipartimenti, atenei), quali che ne siano le finalità. Per stabilire quale relazione debba esistere tra la tipologia (e dimensione) dei soggetti valutati e il peso relativo che nella valutazione dovrebbero avere le due differenti metodologie occorre però preliminarmente esaminare in dettaglio l'origine e il significato dei criteri bibliometrici.⁴

Notiamo che ogni criterio quantitativo trae il proprio fondamento concettuale da considerazioni di tipo statistico, sia che si tratti di un riferimento a valori medi (come nel caso dell'*Impact Factor* o di qualunque altro tipo di *ranking* di riviste o di sedi di pubblicazione) sia che si tratti di valori assoluti individuali (come per il numero delle pubblicazioni e delle citazioni, l'indice *h* e simili) la cui interpretazione dipende comunque dalla comparazione con un grande numero di casi simili, ovvero dalla posizione occupata all'interno di una distribuzione. Ma nelle distribuzioni caratterizzate da una probabilità che, in percentuale, decresce al crescere percentuale del valore del parametro misurato (le cosiddette distribuzioni prive di scala) il significato della media è assai poco pregnante (non potendosi spesso definire la varianza), ed è stato più volte dimostrato, mediante il calcolo delle probabilità, che il rischio che un articolo pubblicato su una rivista di minor rango sia migliore di uno pubblicato su una rivista più "qualificata" non è in genere trascurabile. Sulla base di questa fondamentale premessa è facile argomentare che l'adozione di parametri bibliometrici (in modo esclusivo o prevalente) ai fini di una valutazione comparativa di individui è di per sé un errore concettuale, anche indipendentemente dalla qualità (spesso opinabile) del parametro stesso.

Dovrebbe quindi apparire chiaro che la valutazione comparativa della ricerca individuale, a qualunque fine essa sia effettuata, deve comunque fare necessariamente ricorso a un giudizio di merito sulla produzione scientifica, formulato da valutatori indipendenti e dotati della necessaria competenza disciplinare (*peer review*), che potranno servirsi anche dei parametri bibliometrici interpretandoli alla luce delle proprie esperienze e conoscenze (*informed peer review*). Il riferimento a parametri quantitativi da parte dei revisori, pur non potendo condizionare il giudizio qualitativo, è importante al fine di evitare che elementi di assoluta soggettività possano portare a valutazioni totalmente arbitrarie e a discrepanze gravi tra i giudizi espressi da soggetti differenti.

Nei casi in cui una valutazione, anche non comparativa, abbia comunque l'obiettivo di effettuare una forte selezione, sembra imprescindibile l'intervento di valutatori esperti che siano in grado di associare con competenza al giudizio quantitativo un giudizio di merito qualitativo, che in ultima analisi resta l'unica garanzia di non esclusione di ricercatori di grande valore che per i più svariati motivi (compresa la giovane età, e perfino la reale originalità della ricerca svolta) possano trovarsi a non superare soglie meramente quantitative.

Non è marginale il problema dell'individuazione dei valutatori esperti. Il principio della coerenza con i parametri adottati per la valutazione dei candidati appare generalmente ragionevole, ma ancora una volta è necessaria una specifica attenzione alle differenze anche grandi che certamente esistono fra i differenti settori della ricerca. I rischi sono gli stessi già esposti in precedenza, e ad essi si aggiunge anche quello assai concreto di una restrizione delle possibili rose di "esperti" a numeri così ristretti da costituire vere e proprie caste di professionisti della valutazione, che potrebbero diventare, per banali dinamiche di gruppo, detentori di un potere di scelta e di veto in ultima analisi sproporzionato anche a fronte delle loro pur riconosciute competenze.

In aggiunta a tutte considerazioni di merito fin qui presentate, è doveroso segnalare in questo contesto, anche alla luce di recenti e preoccupanti sentenze dei TAR, il rischio di procedimentalizzazione insito nell'attribuire un valore vincolante ai parametri bibliometrici nella formulazione di giudizi individuali, rischio che ha portato un giurista del calibro di Sabino Cassese⁵ a paventare la possibilità che il giudizio di merito scientifico possa passare "dalle aule universitarie a quelle dei tribunali".

Quando invece il numero dei soggetti, valutati in modo aggregato, è sufficiente a giustificare l'ipotesi che le fluttuazioni di qualità non possano incidere significativamente sull'interpretazione del dato quantitativo, le stesse leggi della statistica ci assicurano un grado di attendibilità dei risultati dell'analisi quantitativa che cresce naturalmente al crescere del numero dei valutati. La comparazione di differenti Dipartimenti relativi alla stessa disciplina, una volta opportunamente standardizzata per tener conto delle dimensioni dei Dipartimenti stessi, può essere quindi plausibilmente effettuata in modo prevalente sulla base di indicatori bibliometrici. Con opportune operazioni di normalizzazione tra discipline differenti, anch'esse possibili su base statistica, l'applicazione dei metodi bibliometrici di valutazione potrebbe costituire anche la base per ripartizioni di risorse almeno parzialmente ancorate al "merito" delle singole strutture all'interno di un Ateneo dato, e anche per la formulazione di un giudizio complessivo sui singoli Atenei.

L'uso degli indicatori bibliometrici per le valutazioni collettive e aggregate non può comunque ridursi a una mera misurazione della quantità di produzione scientifica, che inevitabilmente indurrebbe a comportamenti opportunistici, privilegiando una produzione intensiva, scarsamente meditata e di conseguenza dotata di scarso impatto sulla comunità scientifica nazionale e internazionale. Per questo motivo si è andata sviluppando in alcune comunità di ricerca, soprattutto nelle aree scientifico-tecnologica e biomedica, la ricerca di criteri e di parametri anch'essi di natura quantitativa, ma volti a misurare il grado di interesse con cui le pubblicazioni scientifiche vengono accolte. Malgrado le numerose e anche facili critiche, il conteggio del numero delle citazioni ricevute dagli articoli su rivista sembra generalmente costituire la base di partenza per la stima di questi parametri e indicatori di "impatto scientifico".

A partire dal conteggio delle citazioni individuali si è poi evoluto il concetto di *Impact Factor* (IF) di una rivista, che in sostanza consiste nella valutazione del numero medio annuo delle citazioni per articolo ricevute dagli articoli pubblicati sulla rivista stessa. Come già spiegato più sopra è tuttavia assai improprio convertire l'IF di una rivista in un giudizio sul singolo articolo, e quindi sulla qualità della produzione scientifica individuale. Ma esistono anche fondati motivi, più volte discussi nella letteratura sull'argomento, per ritenere che un'applicazione meccanica dell'IF non produca un'adeguata parametrizzazione della qualità scientifica neanche nel caso di valutazioni collettive, in quanto l'ampiezza dell'intervallo di valori empirici dell'IF (da frazioni dell'unità a diverse decine) è del tutto sproporzionata all'effettiva differenza di qualità media degli articoli pubblicati, e risponde pesantemente a logiche di diffusione editoriale e alle importanti differenze esistenti nelle dimensioni e nelle modalità comunicative delle diverse aree e settori disciplinari, con discrepanze rilevanti anche tra sottogruppi appartenenti allo stesso settore.⁶

Resta comunque confermato che una valutazione “fredda” delle strutture è in linea di principio possibile, e per certi aspetti anche auspicabile (soprattutto in relazione a costi, tempi e sistematicità della valutazione stessa). Ma l’impegno valutativo sarebbe comunque sostanzialmente sterile qualora esso non fosse accompagnato da meccanismi capaci di tradurre le differenze di produttività scientifica così misurate in un effettivo aumento delle risorse per i più meritevoli, erogato a partire dal livello dipartimentale: meccanismi incentivanti o disincentivanti che operino soltanto al livello degli Atenei ben difficilmente otterrebbero effetti positivi sul comportamento delle singole aree e settori scientifici. Una strategia di incentivazione dipartimentale è anche preconditione per una concreta valorizzazione dei nuclei d’eccellenza esistenti anche in quelle realtà che, per ragioni storiche, territoriali e/o organizzative non fossero in grado di raggiungere nel loro complesso i livelli qualitativamente più elevati di produttività scientifica.

Oggetti della valutazione

Per quanto riguarda l’individuazione degli oggetti ai quali applicare criteri e parametri della valutazione, nella realtà italiana attuale occorre sottolineare in primo luogo l’assoluta necessità ed urgenza di consolidare la conoscenza del sistema mediante la creazione di un’attendibile e solida base di dati. Si noti bene che tale compito non è soltanto auspicabile sul piano operativo, ma corrisponde anche a un preciso obbligo giuridico, in quanto la Legge 1/2009, ormai disattesa da quasi quattro anni, prevede l’istituzione dell’Anagrafe Nazionale dei Professori e dei Ricercatori (ANPREPS), nella quale dovrebbe essere inserita tutta la produzione scientifica dei soggetti operanti nel contesto della ricerca pubblica.⁷

Questo strumento, per il quale sono già stati largamente sviluppati gli strumenti tecnici e giuridici, permetterebbe non soltanto una valutazione sistematica di numerosi parametri anche in settori per i quali mancano altre basi di dati adeguate, ma renderebbe anche possibile, con opportuni e già previsti adattamenti, la creazione di un *repository*, ossia di un deposito elettronico contenente non soltanto i metadati (indicazioni bibliografiche) ma anche i veri e propri prodotti della ricerca in formato digitale, con una semplificazione radicale degli aspetti operativi del processo valutativo anche per le procedure volte al giudizio qualitativo che richiedono la concreta disponibilità e accessibilità dei prodotti. Anche i problemi legati alla protezione di questi materiali e alla tutela del *copyright* e della proprietà intellettuale sono tecnicamente risolvibili e non dovrebbero quindi costituire un ostacolo all’avvio e alla realizzazione di un’iniziativa così importante (e tutto sommato non troppo impegnativa).

Il continuo rinvio, con motivazioni largamente incomprensibili, dell’attivazione di questo fondamentale strumento sta facendo mancare a tutti coloro che operano nel campo della valutazione un insieme di dati indispensabile per lo sviluppo delle analisi che sarebbero preliminari a un’adeguata taratura dei parametri quantitativi: sono infatti irrinunciabili, nell’analisi degli indicatori bibliometrici, non solo il calcolo di medie o mediane, ma anche la conoscenza della forma dettagliata delle distribuzioni (anche al fine di individuare eventuali multimodalità in settori non completamente omogenei), la valutazione del grado di proprietà dei prodotti della ricerca nel caso delle collaborazioni (che in taluni settori possono essere anche molto ampie), la conoscenza della dipendenza temporale dei valori medi degli indicatori, il confronto tra i valori nazionali e quelli internazionali (per tutte le discipline per cui tale confronto è già oggi possibile).

Un corollario delle problematiche relative alla costituzione di un’Anagrafe dei prodotti di ricerca è l’identificazione dei prodotti diversi dalle pubblicazioni (quali brevetti, progetti, carte tematiche, etc), che devono certamente trovar posto nella valutazione dell’attività scientifica individuale e aggregata, ma per i quali il problema della classificazione e del riconoscimento di scientificità riveste ancor maggiore difficoltà. Non approfondiremo il tema in questa sede, ma il CUN ha già più volte segnalato l’importanza di un adeguato trattamento della questione.⁸

In stretta connessione con la concreta realizzazione dell'ANPREPS, ma più in generale con qualunque azione di natura valutativa applicata ai prodotti della ricerca, si è aperto negli ultimi anni anche in Italia un ampio dibattito sul tema dell'identificazione della "scientificità" di tali prodotti. La già menzionata Legge 1/2009, anche in questo caso del tutto disattesa, affiderebbe questo compito al CUN, che a sua volta ha già indicato nei propri documenti l'importanza che il giudizio di scientificità sia largamente condiviso all'interno delle singole comunità disciplinari, e ha sottolineato il ruolo cruciale che potrebbe essere svolto dai Nuclei di Valutazione d'Ateneo qualora essi fossero direttamente incaricati di presiedere al controllo sull'immissione delle informazioni relative ai prodotti di ricerca da parte dei singoli ricercatori, anche tenendo conto del fatto che proprio ai Nuclei spetterà in ultima analisi, ai fini della valutazione periodica individuale, il giudizio di congruità delle attività svolte. Soluzioni improvvisate e affidate a gruppi di "esperti", per quanto autorevoli, come quelle adottate negli ultimi tempi sia in relazione all'esercizio di Valutazione della Qualità della Ricerca (VQR 2004-2010), sia in relazione alle procedure per l'Abilitazione Scientifica Nazionale (ASN), oltre a violare il dettato della legge, si espongono troppo facilmente (e si sono in effetti in molti casi esposte) al rischio di superficialità e alle conseguenti legittime contestazioni.

Un punto di partenza per stabilire criteri internazionalmente condivisi è il cosiddetto "Manuale di Frascati" (*Frascati Manual*) redatto per conto dell'OCSE e più volte rivisto fino all'edizione "definitiva" del 2002.⁹ In esso si definisce la ricerca come "lavoro creativo svolto su base sistematica al fine di accrescere la conoscenza, inclusa la conoscenza dell'uomo, della cultura e della società". Sono poi precisate (al punto 64) le nozioni di ricerca di base e applicata e sono elencate (al punto 66) le attività che devono essere escluse dalle nozioni di ricerca e sviluppo, e che sono oggetto di una dettagliata analisi (punti 68-84) al fine precipuo di permetterne la separazione dall'attività scientifica vera e propria. Si rinvia al Manuale stesso per ogni ulteriore approfondimento.

In sostanziale coerenza con i presupposti del Manuale, l'Accademia Norvegese di Scienze e Lettere è giunta a individuare quattro criteri che devono essere simultaneamente soddisfatti affinché una pubblicazione possa definirsi "accademica":¹⁰

1. i risultati presentati hanno caratteri di originalità;
2. i risultati sono presentati in una forma atta alla verifica e/o al riutilizzo in attività di ricerca;
3. la lingua utilizzata e la distribuzione sono tali da rendere la pubblicazione accessibile alla maggior parte dei ricercatori potenzialmente interessati;
4. la sede editoriale (rivista, collana, monografia, sito web) assicura sistematicamente l'esistenza di una *peer review* esterna.

Questi criteri si possono applicare indifferentemente a ogni genere di pubblicazione, e potrebbero essere estesi senza difficoltà anche ad altri prodotti della ricerca. In effetti, come si è già più sopra sottolineato, il riferimento alle "pubblicazioni scientifiche" potrebbe risultare riduttivo e addirittura fuorviante in determinati contesti valutativi se non si desse alla nozione di "pubblicazioni" un'interpretazione estensiva volta a ricomprendere una vasta gamma di prodotti documentabili (quali brevetti, composizioni, disegni, *design*, *performance*, mostre, esposizioni, manufatti, prototipi, opere d'arte e loro progetti, cartografia, banche dati, *software*, e quant'altro).

È anche importante sottolineare che i criteri di scientificità non sono e non possono diventare settoriali: se un prodotto si qualifica come scientifico tale qualificazione deve valere in tutti i contesti per i quali il prodotto stesso risulti pertinente, e tale principio si estende naturalmente alla relativa sede editoriale.

Fatta salva questa specificazione, occorre in primo luogo concentrare l'attenzione sulle possibili definizioni di "scientificità" in riferimento alle forme più tipiche della comunicazione scientifica, in altri termini gli articoli su rivista e le monografie.

Deve risultare innanzitutto chiaro che l'obiettivo di queste definizioni è la formulazione di giudizi comparativi, ma solo l'individuazione di "requisiti minimi", di facile applicazione e verifica, che permettano di analizzare i prodotti sulla base di una logica binaria (accettazione/rigetto). A tal fine sembra inevitabile concentrare l'attenzione non soltanto sul singolo documento ma anche sul contesto e sulle modalità di pubblicazione.

Sulla base delle considerazioni precedenti, l'elemento discriminante per il riconoscimento del carattere di scientificità ad un articolo pubblicato su rivista appare essere l'attribuzione della natura "scientifica" alla rivista stessa. Quest'indicazione appare come una scelta inevitabile alla luce del fatto che è inimmaginabile, sul piano organizzativo, un processo che comporti la valutazione individuale di "scientificità" delle molte decine di migliaia di articoli pubblicati ogni anno.

Per le discipline maggiormente internazionalizzate esistono ormai consolidati repertori di riviste riconosciute come scientifiche, quali ad esempio le basi di dati ISI e SCOPUS, ma ci sono aree e settori che, per differente tradizione culturale o specificità territoriale (quasi sempre legata anche alla lingua in cui sono redatti i testi), non rientrano a tutt'oggi in questa classificazione, anche se numerose riviste di settore sono certamente importanti e autorevoli.

Si tratta quindi di individuare modalità di formazione e verifica del consenso sulla scientificità delle riviste non registrate nei repertori internazionali, tali da permettere alle comunità scientifiche per le quali non esistano ancora in quest'ambito pratiche consolidate e condivise la creazione di propri repertori di "riviste scientifiche" la cui definizione non si basi però sul giudizio di un gruppo ristretto, per quanto autorevole, né tantomeno si fondi sulle indicazioni fornite da società scientifiche di natura e composizione assai disomogenea, per cui non è spesso facile verificarne l'effettiva rappresentatività.

Secondo una concezione della scienza ormai largamente acquisita dalla moderna epistemologia l'appartenenza di un filone di studi e di ricerche a una data disciplina, e quindi la legittimità dei relativi protocolli d'indagine e di esposizione dei risultati, è determinata esclusivamente dal consenso della comunità di coloro che praticano la disciplina stessa, consenso che può evolvere in misura anche molto significativa con il passare del tempo. Da questa premessa sembra conseguire inevitabilmente che l'individuazione di "criteri identificanti il carattere scientifico delle pubblicazioni" è un processo che non può prescindere da un'accurata verifica delle opinioni espresse in materia da ciascuna comunità di studiosi, e che comunque tali criteri devono essere sottoposti periodicamente a una verifica di coerenza con l'evoluzione dei paradigmi e delle modalità di comunicazione delle differenti discipline. A scanso di equivoci si precisa che la definizione della "comunità di studiosi" deve necessariamente avere un orizzonte geografico ampio (possibilmente internazionale) e una dimensione quantitativa tale da allontanare ogni sospetto di autolegittimazione da parte di piccoli gruppi autoreferenziali.

Resta inteso che la validazione della scientificità, quale che sia il percorso adottato a tal fine, non potrà comunque prescindere dal possesso dei criteri sopra elencati, pur dovendosi prevedere una fase di transizione nella quale si dia la possibilità a tutte le comunità scientifiche di uniformarsi all'introduzione di elementi che regolino e in qualche misura certifichino la valutazione ex ante delle pubblicazioni, assicurando anche un tempo sufficiente per l'adeguamento dell'attuale organizzazione editoriale.

Si dovrà garantire nel contempo il riconoscimento della validità della produzione scientifica pregressa, per la quale il criterio della *peer review* non può essere riscontrato oggettivamente e pertanto non può essere assunto retroattivamente come elemento di definizione e di valutazione dei prodotti scientifici.

Ancor più complesso è il tema dell'identificazione del carattere di scientificità quando si passa a considerare le pubblicazioni di natura non periodica. La questione è complicata anche dall'esistenza di alcune tipologie di testi che sono tipiche dell'attività accademica, e che quindi presuppongono competenza scientifica da parte dell'autore, ma che non comportano attività di ricerca originale.

Ancora una volta una scelta di valutazione individuale dei prodotti appare del tutto impraticabile. Non resta quindi che fare riferimento alle modalità di pubblicazione, intendendosi con questo i protocolli di accettazione previsti dalla casa editrice e la collocazione editoriale (all'interno di una collana scientifica, in collana divulgativa o fuori collana). Occorrerà valutare anche le modalità di diffusione dei prodotti editoriali adottate dall'editore (distribuzione in libreria, distribuzione via internet, pubblicazione in rete) per garantire il carattere di accessibilità pubblica, non solo formale, dei lavori scientifici.

Per quanto riguarda poi specificamente gli articoli su rivista dovrà essere valutata con attenzione e con grande cautela l'ipotesi di una classificazione sintetica (e sufficientemente elastica) che identifichi con procedure anche in questo caso ampiamente condivise, all'interno dell'insieme delle riviste ritenute scientifiche, differenti livelli di qualificazione, garantendo comunque un'equilibrata ripartizione delle riviste tra i suddetti livelli. Il "punteggio" qualitativo da attribuire ai diversi livelli non dovrebbe in ogni caso essere riferito a un intervallo di valori troppo ampio.

La determinazione di questo *pool* di riviste dovrebbe comunque risultare da un processo interattivo, nel quale le comunità scientifiche, anche tramite le proprie associazioni, giochino un ruolo determinante, e nel quale i criteri di riferimento, oltre quelli già in precedenza definiti in relazione alla scientificità, siano l'impatto (nazionale e meglio ancora internazionale) della rivista, la presenza nelle principali banche dati internazionali, l'autorevolezza della direzione scientifica, l'affidabilità della gestione organizzativa.

Per quanto riguarda le monografie, il cui ruolo nella trasmissione dei risultati della ricerca è particolarmente importante nel campo delle scienze umane, sembra molto arduo immaginare criteri di classificazione "oggettiva" dell'impatto scientifico che si basino sulla collocazione editoriale o su un qualche tipo di *ranking* delle Case Editrici, una pratica peraltro di dubbia legalità e del tutto sconosciuta in tutti i Paesi che hanno affrontato il problema della valutazione della produzione scientifica. Qualora si intenda, in sede valutativa, graduare in qualche misura il valore delle opere pubblicate in volume sembra quindi ineludibile stabilire, anche nel caso di valutazione aggregata, una qualche forma di giudizio comparativo mediante *peer reviewing*, malgrado l'impegno straordinariamente più elevato insito in tale procedura

Differenze strutturali tra le aree

Non soltanto la valutazione individuale, ma anche quella aggregata si scontra inevitabilmente con l'esigenza di tener conto di differenze di natura fondamentalmente qualitativa, soprattutto quando entra in gioco l'esigenza di effettuare confronti tra aree scientifiche anche molto diverse, anche ai fini di un'adeguata ripartizione delle risorse. Un contributo in tal senso potrà comunque venire da un approfondimento degli studi volti alla comparazione delle distribuzioni dei valori dei principali indicatori bibliometrici. Esistono già significativi risultati che evidenziano il carattere di universalità di tali distribuzioni, quando opportunamente normalizzate, in particolare per quanto riguarda le citazioni e l'indice di Hirsch.

Anche un confronto tra l'esito della VTR (2001-2003) e il risultato di un'esplorazione puramente quantitativa (e dichiaratamente rozza) dei dati disponibili in rete, effettuata con l'uso di *Google Scholar*, sembra indicare che, quando presi in esame con le dovute cautele, anche i dati bibliometrici potrebbero consentire un confronto abbastanza attendibile tra le differenti aree scientifiche.

Nel campo delle scienze umane e sociali la situazione è resa ancor più complessa dall'assenza di basi di dati sufficientemente ampie e adeguate per il conteggio delle citazioni, e dal ruolo rilevante della produzione monografica, per la quale tale conteggio è intrinsecamente più difficile, se non addirittura impossibile, almeno con gli attuali strumenti di analisi.

In ogni caso la definizione di criteri bibliometrici per la valutazione della ricerca richiede un ripensamento complessivo, con l'auspicabile obiettivo di identificare, anche per le scienze umane, criteri che da un lato siano per quanto possibile omogenei tra le differenti discipline, e dall'altro non penalizzino artificialmente differenti tradizioni culturali e modalità di comunicazione scientifica.

Conclusioni

Le soluzioni proposte fino a questo punto sono comunque da considerarsi largamente contingenti. Se infatti proviamo a collocarci in una prospettiva temporale più ampia possiamo facilmente immaginare che nel medio periodo (ovvero nei prossimi 5-10 anni) l'evoluzione delle modalità di comunicazione, anche scientifica, associata alla diffusione e allo sviluppo di Internet e dell'editoria elettronica, e soprattutto delle moderne forme di interazione collettiva (*social network* e simili) identificate anche dalla locuzione *Web 2.0*, porterà a nuovi meccanismi valutativi, al momento ancora in embrione nel campo della ricerca, ma già efficaci in altri campi di attività. Ci riferiamo ad esempio all'idea del "controllo aperto" lanciata da *Nature*, e consistente nell'idea di "esporre" i manoscritti per un certo tempo su Internet, con la possibilità di essere commentati da tutti i lettori interessati, prima che il comitato editoriale decida di "pubblicarli" (probabilmente soltanto in forma elettronica), anche sulla base dei giudizi che sono stati espressi. Possiamo anche aspettarci sostanziali raffinamenti di quella forma di controllo *ex post* che è data dalle citazioni, se si supererà l'attuale meccanismo centralizzato (fondato quasi esclusivamente su basi di dati proprietarie) in favore di un'analisi a tutto campo mediante motori di ricerca (sul modello, per quanto certamente ancora molto rozzo, proposto da *Google Scholar*), che permetterebbe facilmente di estendere questo tipo di valutazione anche alle monografie, soprattutto se le forme di pubblicazione elettronica, come prevedibile, andranno generalizzandosi, auspicabilmente con modalità *open access*.

In quest'ottica è più che mai importante mantenere un atteggiamento aperto nei confronti dei meccanismi e delle regole per la valutazione, evitando irrigidimenti normativi e burocratici che rischierebbero di risultare superati in un breve arco di tempo, e quindi di frenare pericolosamente il pieno dispiegarsi delle potenzialità offerte alla ricerca realmente creativa dalle nuove forme della comunicazione scientifica.

-
- ¹ AA.VV., *Malata e denigrata*, a cura di M. Regini, Roma, Donzelli 2009
- ² P. Rossi, *La valutazione della ricerca*, *Analysis* Anno 12 – N.2/2010, pp.4-7
- ³ P. Rossi, *Problemi e ipotesi per valutare la ricerca*, *Menodizero*, Anno II, N.5, Aprile-Giugno 2011
- ⁴ A. Baccini, *Valutare la ricerca scientifica. Uso e abuso degli strumenti bibliometrici*, Bologna, Il Mulino 2010
- ⁵ S. Cassese, *L'ANVUR ha ucciso la valutazione. Viva la valutazione!*, <http://www.roars.it/online/lanvur-ha-ucciso-la-valutazione-viva-la-valutazione/>
- ⁶ A. Figà Talamanca, *L'Impact Factor nella valutazione della ricerca e nello sviluppo dell'editoria scientifica*, comunicazione al IV seminario del Sistema Informativo Nazionale per la Matematica, Lecce 2 ottobre 2000, <http://www.roars.it/online/limpact-factor-nella-valutazione-della-ricerca-e-nello-sviluppo-delleditoria-scientifica/>
- ⁷ Legge 9 gennaio 2009, n.1 *Conversione in legge, con modificazioni, del D.L. 10 novembre 2008, n.180*, G.U. n.6 del 9/1/2009
- ⁸ Dossier CUN n.3 *"Anagrafe delle ricerche e valutazione"* (a cura di A. Lenzi, P.Rossi e N. Pistoia), Settembre 2012
- ⁹ OECD, *Frascati Manual: Proposed Standard Practice for Surveys on Research and Experimental Development*, 6th edition (2002)
- ¹⁰ NAHEI committee, *A Bibliometric Model for Performance-based Budgeting of Research Institutions*