

Paola Galimberti

Valutarsi e aprirsi: alternative per la valutazione

Scopo di questo intervento è quello di verificare se e quanto l'attuale sistema di valutazione implementato da ANVUR abbia inciso sui comportamenti e le scelte editoriali degli studiosi di ambito umanistico a partire dai dati di un grande ateneo multidisciplinare del Nord-Italia. Sulla base delle analisi fatte si tenterà di avanzare delle proposte alternative per la valutazione nelle scienze umane maggiormente rispettose delle modalità attraverso le quali gli studiosi, in particolare dell'Area 10, sono soliti comunicare i risultati della loro attività scientifica.

Partiamo dai dati

Lavoro coi dati, nella convinzione che i dati, se affidabili e letti nel contesto da chi sa interpretarli, possano essere un eccezionale supporto informativo per le decisioni della *governance* di un ateneo.

Poiché l'Italia non ha ancora trovato il modo di rendere disponibili i dati sulla ricerca (pubblicazioni, progetti) a livello nazionale (la anagrafe nazionale della ricerca prevista dalla legge 1/2009 non è mai stata realizzata), ho utilizzato gli unici dati per me certi, che sono quelli dell'Università degli Studi di Milano che, come molti altri atenei italiani, dispone di un archivio istituzionale (IRIS) il cui contenuto è però certificato e dunque utilizzabile per analisi e monitoraggio.

Un primo elemento di indagine è stato quello di capire quali fossero le tipologie di lavori presentati per l'esercizio di valutazione nazionale (VQR 2011-2014) dalle aree CUN relative alle scienze umane e sociali per verificare quali tipologie sono state considerate esprimere al meglio la qualità della ricerca nelle aree 10-14.

Il risultato della analisi è rilevabile dalla figura 1:

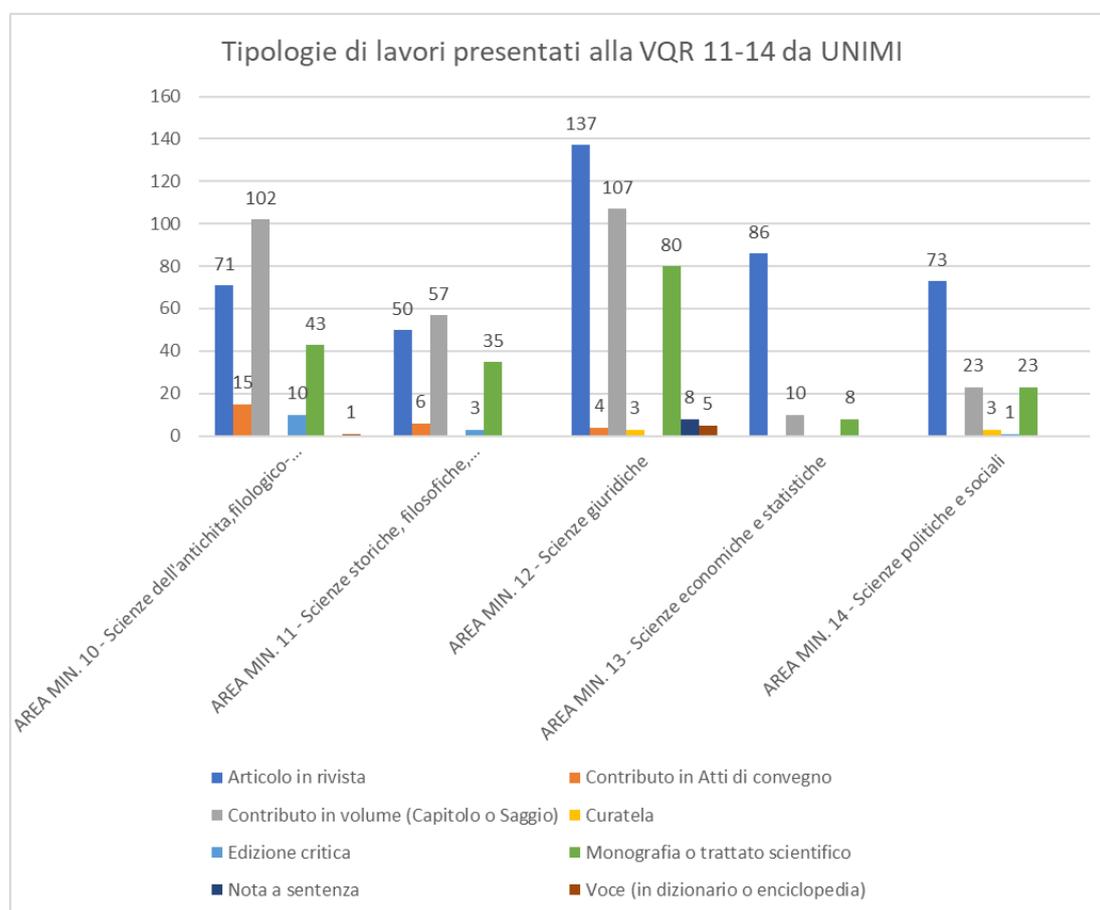


Fig. 1 Tipologie di lavori presentati nelle aree 10-14 per la VQR 11-14

Dalla figura 1 emerge chiaramente come le diverse aree si siano orientate diversamente rispetto alle tipologie di pubblicazioni selezionate, con le aree 10 e 11 che hanno privilegiato gli interventi in volume, le aree 13 e 14 che hanno privilegiato gli articoli e l'Area 12 che si trova a cavallo con un numero di contributi in volume e monografie consistente, ma anche con un buon numero di articoli su rivista.

Se consideriamo le stesse tipologie selezionate per la VQR11-14 distribuite sugli 11 dipartimenti di area umanistica dell'Università di Milano otteniamo lo stesso risultato (fig. 2), con i dipartimenti di Area 10 e 11 per i quali gli articoli non sembrano affatto esprimere il meglio della ricerca degli afferenti al Dipartimento.

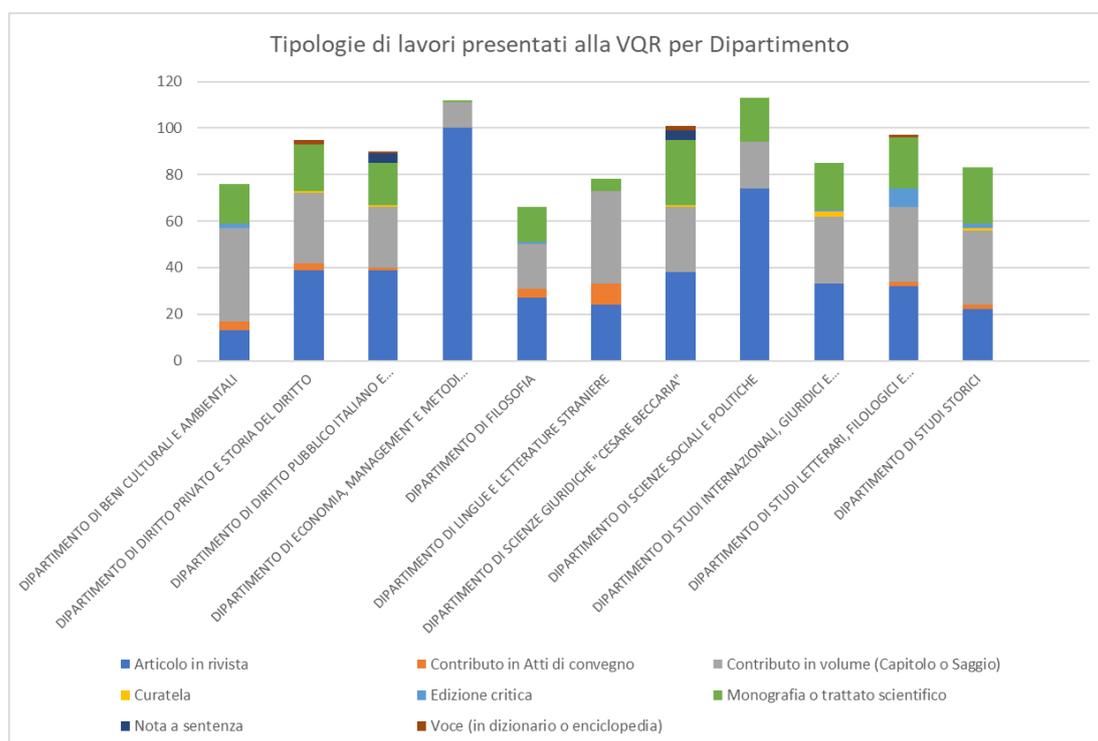


Fig. 2 Tipologie di lavori presentati alla VQR11-14 distribuite sui dipartimenti

Ma cosa è successo dopo la VQR? Come si sono orientate le scelte dei ricercatori ad esempio di Area 10?

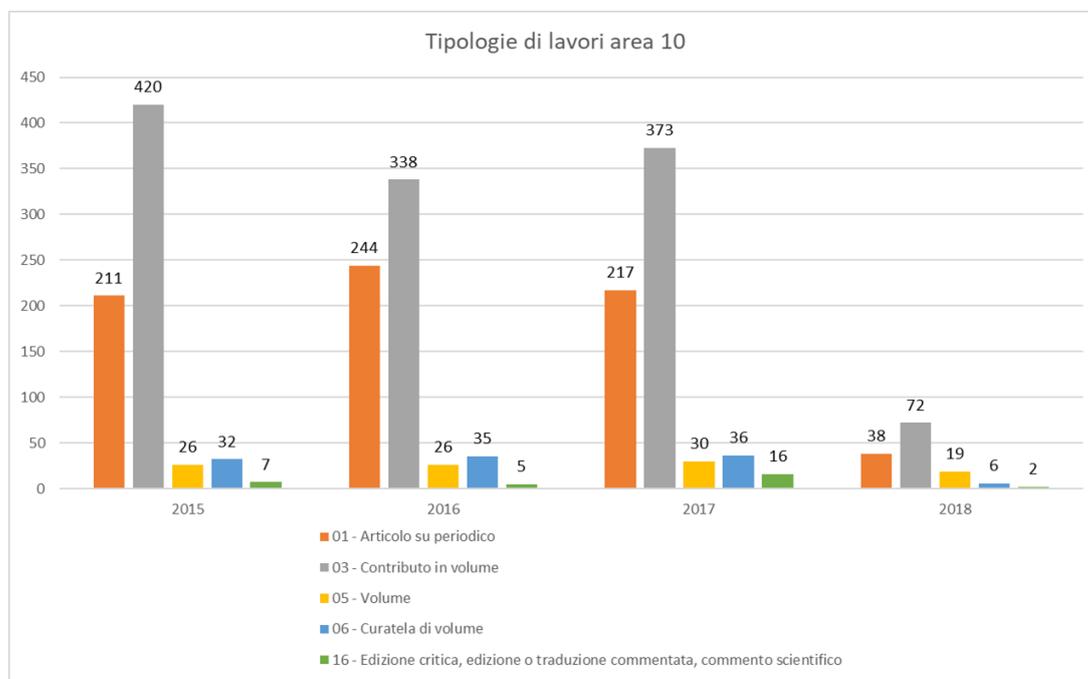


Fig. 3 Tipologie di lavori di ricerca per l'Area 10 nel periodo 15-18

Come si evince dalla figura 3 anche dopo la VQR nel periodo 2015-2017 (per il 2018 i dati risultano ovviamente incompleti) in Area 10 la tipologia di lavori più frequente resta il contributo in volume con leggere flessioni in più o in meno nel corso degli anni.

Il senso di questa analisi è quello di capire se davvero per l'Area 10 sia possibile porre l'enfasi sulle riviste e in particolare sulle riviste di fascia A e se la regola di considerare i risultati della VQR per la promozione o il mantenimento di una rivista nella fascia A possa essere applicata in maniera significativa per quest'area. Ricordiamo che la pubblicazione in riviste di fascia A incide su molti esercizi: la Abilitazione scientifica nazionale, l'accreditamento dei dottorati di ricerca. È altrettanto vero che in maniera distorta la promozione o il mantenimento nella fascia A dipende da numero e valutazione delle pubblicazioni in fascia A presentate per la VQR.

Se teniamo conto che la maggior parte dei lavori presentati alla VQR (fig. 1 e 2) e anche di quelli successivi alla VQR (fig. 3) non appartiene alla tipologia articolo, se cioè l'articolo non è tipologia particolarmente frequentata dall'Area 10, ci si chiede sulla base di quale considerazione il numero e la valutazione degli articoli presentati per la VQR possa essere discriminante per la promozione in fascia A di una rivista o per la sua eventuale bocciatura.

Quali effetti ha avuto la valutazione in Italia?

Dopo due campagne di valutazione nazionale e dopo due tornate di Abilitazione scientifica Nazionale è forse lecito chiedersi quali siano stati gli effetti sul sistema, e cioè se l'introduzione della valutazione ed in particolare di una valutazione quantitativa massiva abbia avuto un effetto positivo sulla ricerca. Le campagne di valutazione sia a livello di istituzioni, sia e soprattutto a livello individuale hanno avuto un effetto sulle modalità di produzione scientifica dei ricercatori? E se sì l'auspicabile miglioramento connesso ad esercizi di valutazione quantitativa massivi è avvenuto?

Può essere utile ampliare la analisi alle aree STEM, cioè a quelle aree definite da ANVUR *bibliometriche*¹.

Dalla figura 4 si evince chiaramente quale sia stato l'effetto della legge Gelmini e in particolare della operatività dell'ANVUR, un fenomeno che appare con ancora maggiore evidenza, soprattutto per i ruoli soggetti a progressione (associati e ricercatori):

¹ Usiamo qui la definizione di ANVUR precisando che la bibliometria può essere applicata a qualsiasi area per cui la definizione risulta imprecisa.

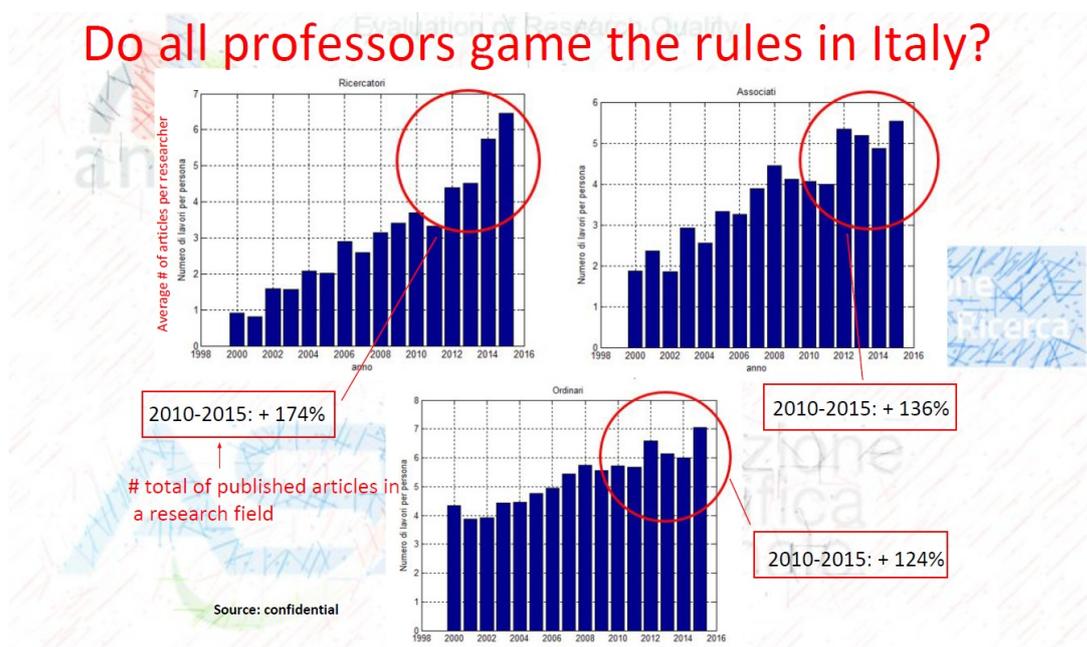


Fig. 4 tratta da A. Baccini, *Performance based incentives, research evaluations and the percolation of bad science*, Berlin, May 2018

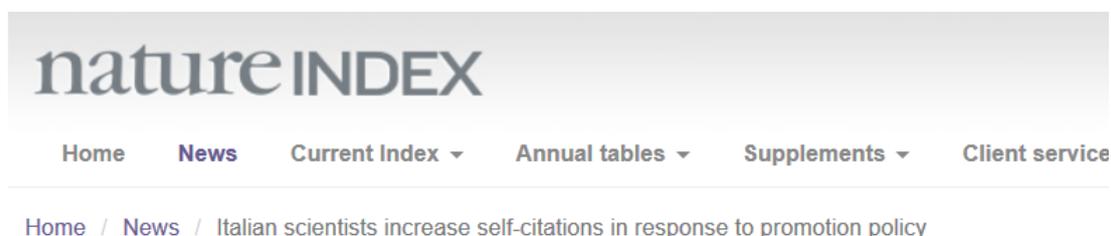
Certamente si assiste ad un fenomeno di aumento delle produttività, anche se ancora non siamo in grado di dire come l'aumento del numero di pubblicazioni scientifiche si traduca in miglioramento della loro qualità.

Introduciamo a questo punto una frase tratta dall'ultimo libro di Andrea Bonaccorsi, apparso per Springer, e che raccoglie in larga parte gli esiti del Primo Concorso pubblico di idee di ricerca per la valutazione nelle scienze umane lanciato nel 2014².

*In STEM, citations have **unambiguous meaning** of credit assigned to authors that made the previous discoveries* (Bonaccorsi 2018)

La affermazione, a dire il vero all'epoca dei sistemi di finanziamento della ricerca *Performance based* alquanto ardita, appare contraddetta da una recente ricerca di cui troviamo notizia su *Nature index* (Seeber *et al.* 2019):

² Di questo bando è uscita di recente la terza tornata.



[Share on Facebook](#) [Tweet this article](#)

Italian scientists increase self-citations in response to promotion policy

Study reveals how research evaluations can lead to self-serving behaviour.

Secondo questa ricerca gli scienziati italiani avrebbero risposto in maniera adattativa al sistema di valutazione per le progressioni di carriera, incrementando il numero delle autocitazioni. Quindi nelle aree cosiddette bibliometriche si assiste a fenomeni opportunistici sia rispetto alla produttività che rispetto alla “qualità” misurata in numero di citazioni.

È interessante vedere se questi fenomeni opportunistici hanno preso piede allo stesso modo in tutte le aree e per capirlo si è presa in considerazione la produttività dal 2004 ad oggi di alcuni ricercatori e professori di Area 10 dell’ateneo milanese.

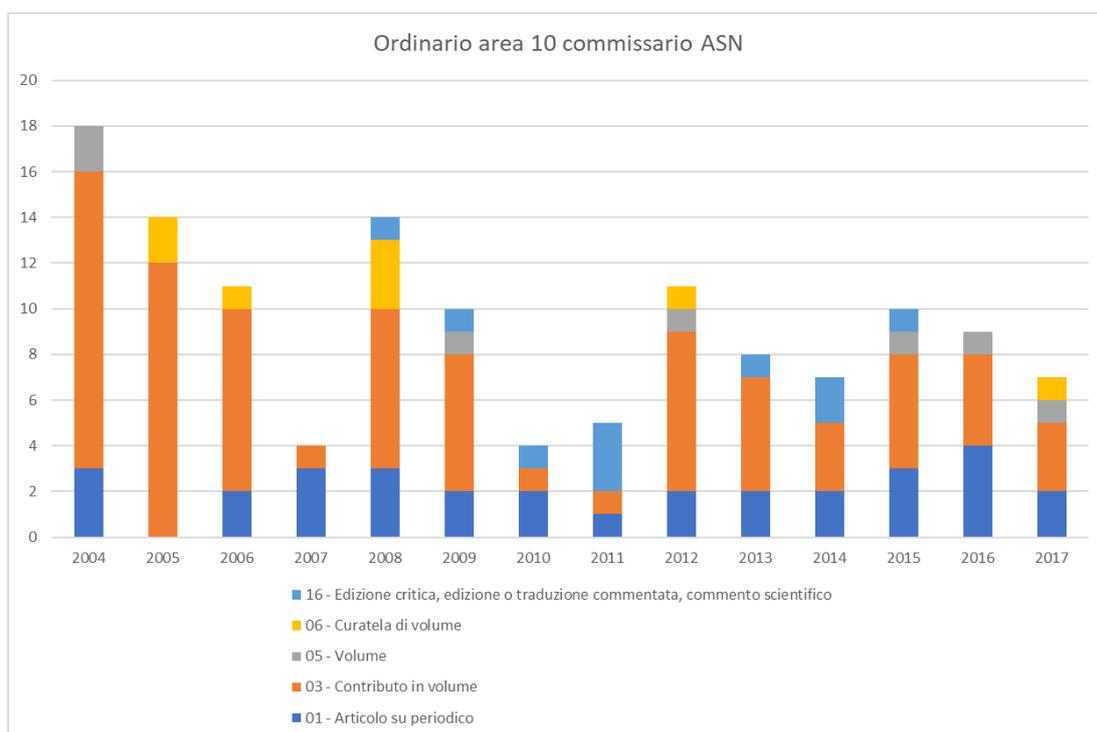


Fig. 5 numero e tipologia di pubblicazioni di ordinario commissario ASN

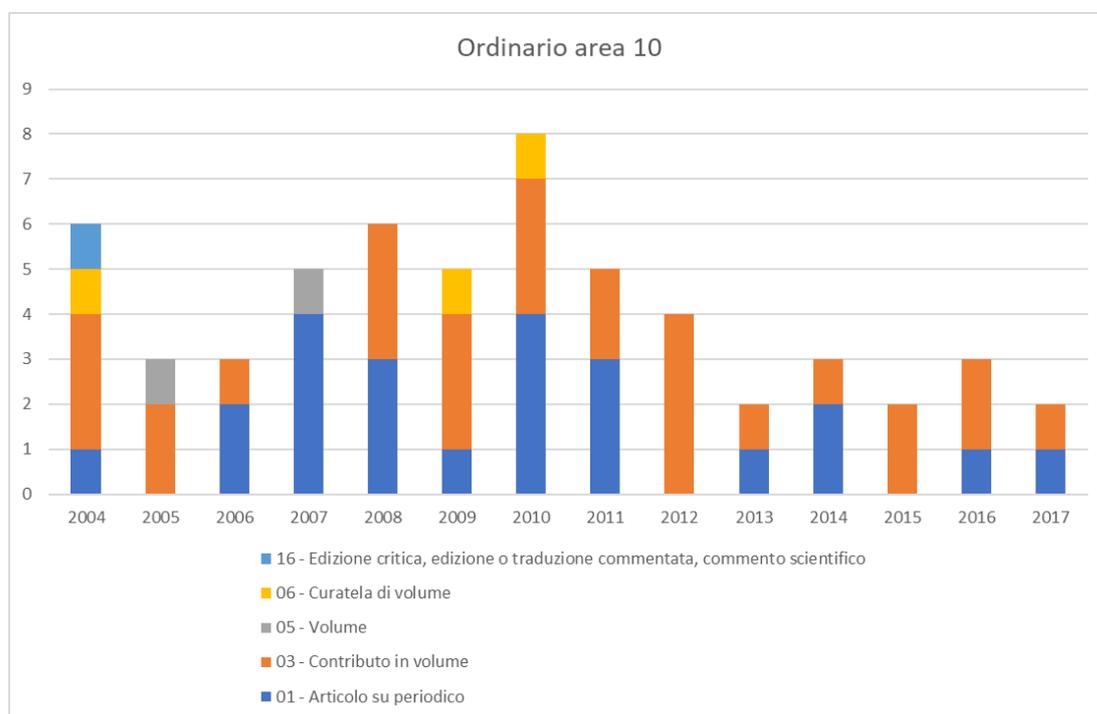


Fig. 6 numero e tipologia di pubblicazioni di ordinario

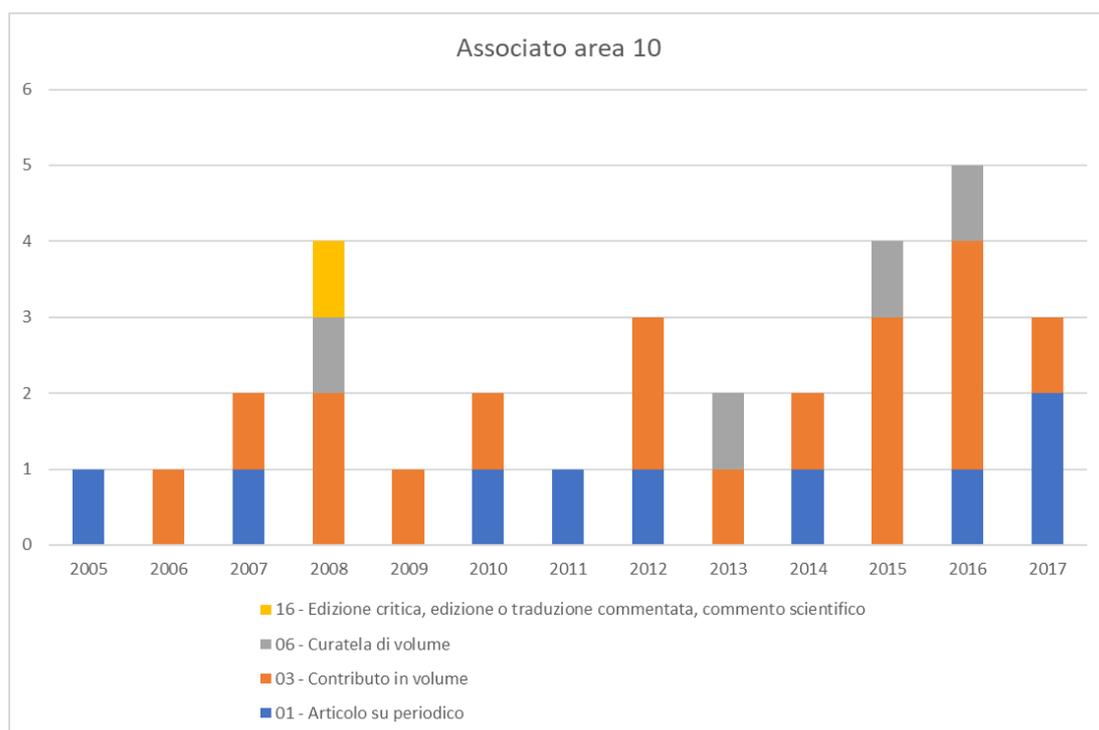


Fig. 7 numero e tipologia di pubblicazioni di associato

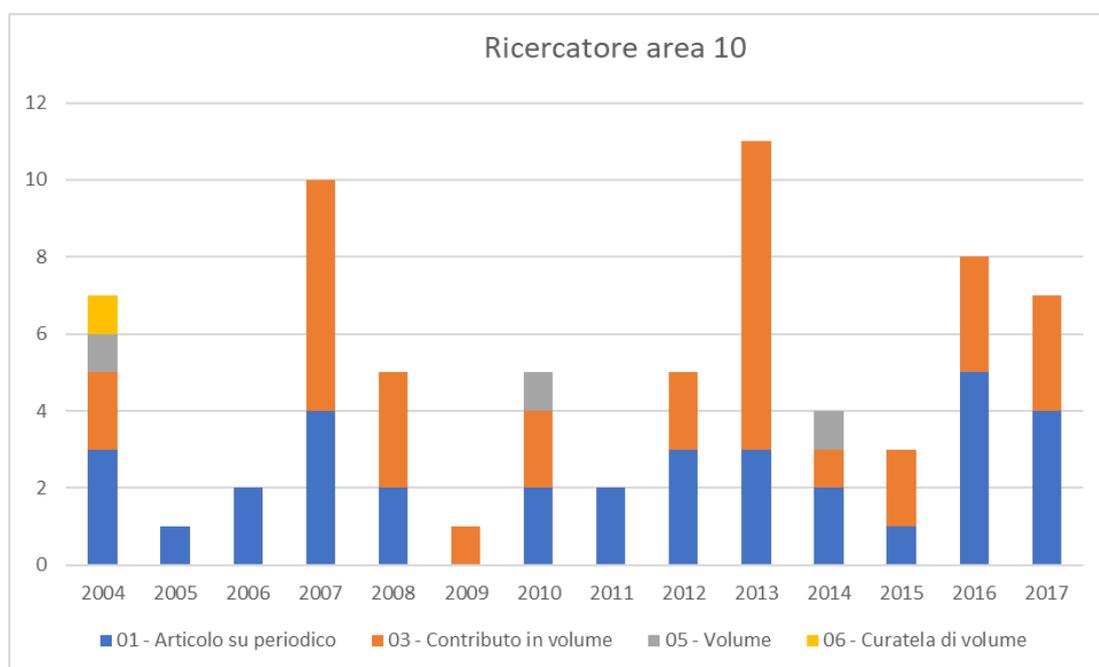


Fig. 8 numero e tipologia di pubblicazioni di ricercatore

Come si può vedere dai grafici la situazione in tutti i ruoli non è così ben delineata come ci si aspetterebbe, sia per quanto riguarda la produttività che non è costante nel tempo e descrive la tipica alternanza fra periodi di ricerca e periodi di scrittura, sia per quanto riguarda la scelta delle tipologie.

I dati ci spingono dunque a dire che nelle scienze umane (o per lo meno in Area 10) le strategie di pubblicazione continuano ad essere diversificate. L'enfasi posta negli ultimi esercizi di valutazione sulle riviste di fascia A non trova giustificazione a nessuno dei livelli e dei ruoli presi in considerazione. Gli articoli su rivista (di qualsiasi tipo) continuano a non rappresentare il canale privilegiato di comunicazione. Appare quindi poco opportuna e soprattutto poco significativa la costruzione di indicatori che si basano sulle riviste di fascia A: collegi di dottorato, valutazione delle riviste sulla base degli articoli (numero e voto ottenuto) presentati per la VQR, soglie ASN.

A maggior ragione appare ambigua la seguente affermazione di Andrea Graziosi, ex Presidente dell'ANVUR e grande sostenitore della necessità dei *ranking* di riviste che nel 2017 ha incontrato la delegazione di Area 10 «A margine Andrea Graziosi ha fatto presente che nella seconda VQR il numero degli articoli su rivista sottoposto a valutazione è **umentato significativamente**» (2017 ANVUR incontra la delegazione di Area 10, Resoconto). Cosa intendeva dire il Presidente dell'ANVUR? Che l'effetto adattativo ha coinvolto anche l'Area 10? Che la valutazione di Stato condiziona pesantemente gli stili di pubblicazione in modo da poter controllare meglio ciò che si pubblica?

Non discuteremo il tema delle riviste di Fascia A

Non è questa la sede per approfondire la tematica delle riviste di fascia A, tematica su cui è stato scritto moltissimo³, e tuttavia si fa presente che sono molti i canali di pubblicazione prestigiosi, su cui pubblicano i nostri ricercatori e che, pur essendo indicizzati in Scopus o in WOS, non trovano posto nell'elenco italiano. Nella sola Area 10 e nell'archivio istituzionale dell'Università di Milano se ne contano almeno 20. Vale a dire 20 sedi di pubblicazione che hanno superato un complicato meccanismo di valutazione per essere indicizzate e che non sono spendibili dagli autori per nessuno degli esercizi nazionali.

Non approfondiremo dunque il tema delle liste di riviste di fascia A, ma ci concentreremo invece sulle caratteristiche che una lista di autorità (quale la lista delle riviste di fascia A aspira ad essere) dovrebbe avere:

³ Si veda qui ad esempio: <https://www.roars.it/online/tag/riviste-di-fascia-a/>.

In primo luogo una simile lista elenca tutte le riviste (o eventualmente le collane) che rispondono a **criteri formali di qualità e trasparenza validi in tutto il mondo, verificabili da chiunque e accertate da specialisti dell'informazione**⁴

Come ben sottolineano Bon, Taylor e McDowell la valutazione sulla base del valore del contenitore (e non del contenuto) lascia a coloro (pochi) che stanno nei *board* delle riviste il diritto di vita o di morte rispetto a coloro (molti) che devono affrontare progressioni di carriera.

«Journal based evaluation creates an asymmetry within the scientific community between a minority of scientists sitting in editorial boards, who have the exclusive power to give value to a scientific article by accepting it in their journal» (Bon, Taylor, McDowell, 2017).

Ovviamente lo stesso meccanismo perverso si innesca quando quegli stessi professori **che stanno nei board** delle riviste siedono anche **nell'organo che deve decretare** il loro ingresso o la loro espulsione da una lista di autorità.

La assenza di conflitto di interessi è dunque un punto fondamentale nella costruzione dei gruppi di esperti (non necessariamente professori) che devono lavorare sulla lista di autorità.

Un altro requisito fondamentale di una lista di autorità che deve essere utilizzata a scopi valutativi è che viene aggiornata in tempo reale ogni volta che un ricercatore pubblica in un nuovo canale (tenendo sempre presenti le indicizzazioni pregresse: DOAJ, WOS, ERIH, Scopus). Ritardi nella indicizzazione possono causare un danno alla carriera delle persone e non è dunque possibile prevedere aggiornamenti a lungo termine o solo in determinati periodi dell'anno.⁵

Se queste caratteristiche sono garantite, allora solo gli articoli pubblicati in riviste presenti nella lista di autorità sono presi in considerazione negli esercizi di valutazione, fermo restando che la valutazione non può e non deve dipendere dal contenitore.

Ma quali alternative possono esserci ad una valutazione quantitativa?

Siamo in un'epoca in cui ai ricercatori si aprono numerose possibilità per trasmettere alla propria comunità scientifica ma anche al mondo le proprie ricerche. Il mezzo digitale rappresenta una opportunità unica per le scienze umane di raggiungere colleghi in ogni parte del mondo, e la apertura dei contenuti permette l'avvio di un dibattito all'interno delle comunità scientifiche. Gli archivi istituzionali o disciplinari,

⁴ Il modello potrebbe essere DOAJ o anche ERIH nel mondo *open*, oppure Scopus o WOS nel mondo commerciale.

⁵ Ricordiamo che la lista di Area 13 è stata pubblicata a febbraio (come le altre) ma è uguale a quella precedente. Non c'è dunque in questa lista alcun nuovo ingresso.

gli archivi di *pre-print*, le riviste *open access* che nelle scienze umane sono spesso senza costi per chi legge e senza costi per chi pubblica, i *megajournals* o i portali disciplinari sono solo alcune delle manifestazioni⁶ a cui ha condotto l'avvento del web.

In questo nuovo contesto solo le modalità quantitative con cui la ricerca viene valutata sembrano essere rimaste immutate, anzi, il tentativo che viene portato avanti è quello di espandere sistemi che già si sono rivelati fallimentari per le aree STEM anche alle scienze umane e sociali.

La proposta è dunque quella di riportare la bibliometria all'interno del contesto in cui è nata che è quello della descrizione della scienza. In tutte le situazioni in cui i finanziamenti vengono distribuiti in maniera automatica sulla base di indicatori si sono innescati meccanismi adattativi (*publish or perish*), minando la credibilità della scienza e degli studiosi.

Se i ricercatori orientano le loro scelte (ambito di ricerca, sede di pubblicazione e tipologia) sulla base degli incentivi è necessario separare il momento della valutazione da quello della distribuzione automatica di fondi.

Le scienze umane hanno fino ad ora mostrato una adesione moderata alla rivoluzione digitale e soprattutto al movimento dell'accesso aperto. A questo punto però, soprattutto nelle aree che sono state meno toccate dai sistemi di valutazione quantitativa si presenta l'occasione di intraprendere percorsi diversi prima di essere investite dalla fiducia nel numero.

In accordo con quanto la Commissione Europea prevede per la ricerca che finanzia è importante che siano aperti e trasparenti tutti i processi attraverso i quali la ricerca viene validata (*peer review*) disseminata (pubblicazione) e valutata. La apertura dei contenuti consente il confronto con le altre comunità disciplinari, favorisce l'interdisciplinarietà che in questo momento è mortificata e appiattita su generi e canali di comunicazione definiti da ristretti gruppi di potere (si pensi a fatto che le riviste di fascia A sono tali solo per determinati settori concorsuali). La apertura dei contenuti permette inoltre il confronto con chi sta al di fuori del sistema innescando certamente un circolo virtuoso (*citizen science*).

In quest'ottica anche la *peer review* perde il suo carattere segreto e confidenziale (decisione binaria, dentro o fuori, da parte di due o tre esperti) e si trasforma in discussione, dibattito pubblico all'interno della comunità scientifica volto al raggiungimento di un consenso diffuso sul valore di uno studio⁷. La accessibilità dei *report* di valutazione di un lavoro farà in modo che essi siano accurati, che i conflitti di interesse siano dichiarati e visibili a tutti, il giudizio sia motivato e la decisione di

⁶ *Open edition* <https://www.openedition.org/>, *PKP* <https://pkp.sfu.ca/ojs/>, *Science open* <https://www.scienceopen.com/>, *Open library for the Humanities* <https://www.openlibhums.org/> sono solo alcuni esempi.

⁷ Alcuni esempi potrebbero essere *F1000 research* o il *Self journals of Science*.

pubblicare un lavoro sia presa in maniera trasparente e verificabile da chiunque abbia interesse per un determinato argomento.

Anche i dati utilizzati per la valutazione devono essere disponibili a tutti insieme ai *report* di valutazione per permettere ad altri ricercatori di fare gli stessi test validando così i risultati⁸. Ricordiamo che i *dataset* utilizzati da ANVUR per gli articoli che danno giustificazione scientifica dei propri sistemi di valutazione non sono disponibili e che quindi la robustezza dei metodi utilizzati può essere certificata oggi solo dall'ente che li utilizza.

Il tema della apertura non riguarda solo il momento della valutazione, ma tutto il processo di produzione e validazione della scienza, in qualsiasi ambito. Ciò che deve avvenire non è quindi solo la modifica di un regolamento o di un bando o di un criterio o di un indicatore, ma un vero e proprio mutamento culturale. Da un lato questo cambiamento è facilitato nelle scienze umane dalla mancanza di modelli prestabiliti, dall'altro risulta frenato da una sorta di ritegno ad esporsi tipico degli umanisti, da una specie di gelosia per i propri testi e da un legame ancora molto forte con l'editoria tradizionale (soprattutto nazionale) che fatica ad adeguarsi al mutato contesto.

È un cambiamento che tuttavia non deve avvenire all'improvviso, oggi i lavori anche degli umanisti reperibili ad accesso aperto sono molto aumentati grazie al supporto delle politiche locali ed europee e nonostante la totale assenza di politiche ministeriali al riguardo⁹.

Ci sono tuttavia segnali positivi che fanno ben sperare nella capacità delle scienze umane di intraprendere il cambiamento. Iniziative come UCL press (l'*university press* di UCL che pubblica soprattutto lavori nell'ambito delle scienze umane e sociali) ha festeggiato da poco il milione di download. Una piattaforma come quella delle riviste dell'Università di Milano¹⁰ registra oltre 700mila download l'anno, da ogni parte del mondo. La strada imboccata (comunicazione scientifica digitale e aperta) sembra quella più corretta per la diffusione della cultura umanistica e per una sua valutazione trasparente, pare dunque importante per le scienze umane percorrerla con convinzione, evitando di scimmiettare comportamenti e modalità tipici di altre aree scientifiche che hanno ormai rivelato tutta la loro fallibilità.

⁸ Si rimanda qui al concetto di *FAIR data*. I dati della ricerca devono essere *findable, accessible, interoperable e reusable*.

⁹ Di apertura delle pubblicazioni o dei dati si parla nei PRIN ad esempio, ma senza che vi sia un controllo ex post di quanto fatto dai ricercatori. L'Italia si caratterizza in ambito europeo per la totale assenza di politiche sulla apertura, cosa che rende più difficile una loro implementazione a livello locale. Gli ottimi risultati ottenuti da alcune istituzioni sono davvero un successo.

¹⁰ <https://riviste.unimi.it/>.

Riferimenti bibliografici:

BACCINI 2018

A. Baccini, *Performance Based Incentives, Research Evaluations and the Percolation of Bad Science*, relazione tenuta a: G20-Global solutions summit, Berlin, 28-29 May 2018.

BONACCORSI 2018

A. Bonaccorsi, *The Evaluation of Research in Social Sciences and Humanities: Lessons from the Italian Experience*, Springer 2018.

BONN – TAYLOR – MC DOWELL 2017

M. Bonn, M. Taylor, G.S. Mc Dowell, *Novel Processes and Metrics for a Scientific Evaluation Rooted in the Principles of Science*, «Self journals of science» 2017.

<http://www.sjscience.org/article?id=580%20>.

FRANZONI – SCELLATO – STEPHAN 2011

C. Franzoni, G. Scellato, P. Stephan, *Changing Incentives to Publish*, «Science» CCCIII, 702-703.

SEEBER ET AL. 2019

M. Seeber, M. Cattaneo, M. Meoli, P. Malighetti, *Self Citation as Strategic Response to the Use of Metrics for Career Decisions*, «Research Policy» XLVIII.2, 478-91.