



Alla ricerca del Santo Graal (e come vivere senza)

2014-06-16 09:44:22 By Paolo Pini

L'accademia italiana sembra impegnata in una corsa alla ricerca del Santo Graal dell'Indicatore Perfetto di Qualità (IPQ) per valutare la ricerca scientifica. Man mano che i Cavalieri si addentrano in terre sconosciute ed ostili, si annuncia la scoperta di indicatori sempre più vicini all'IPQ: il numero di pubblicazioni, le loro citazioni, la VQR individuale, l'Impact Factor, la peer review, e decine di altri indicatori. Meglio un approssimato (o approssimativo) Graal che non averne nessuno! Come si può vivere senza la certezza che il giusto sarà distinto dal peccatore, l'inattivo dall'eccellente? Il consenso raccolto dalla necessità di valutare la ricerca è talmente plebiscitario da meritare un po' di sano scetticismo illuminista. Che la valutazione della ricerca sia, o rischi di divenire, una leggenda priva di fondamento?



THE LAST APPEARANCE OF THE SANGREAL



L'accademia italiana in tutte le sue articolazioni istituzionali (strutture, commissioni, nuclei di valutazione, ecc.) sembra impegnata in una frenetica corsa alla ricerca del Santo Graal dell'Indicatore Perfetto di Qualità (IPQ) da usare per valutare il lavoro di chiunque svolga attività di ricerca scientifica.

L'idea di fondo che sembra universalmente condivisa è che ogni ricercatore sia dotato di una sua precisa, universale, oggettiva misura di qualità, come una sorta di anima, in base alla quale, in un mondo perfetto, andrebbe remunerato, promosso, premiato, e con lui chi lo assume, etc. con altrettanta giustizia di quanto l'Altissimo applicherà nel giorno del Giudizio. Peccato però che, per noi mortali, questa misura sia non osservabile, e quindi siamo costretti ad accontentarci di misure indirette, sicuramente imprecise, che (si spera) siano fortemente correlate con l'IPQ. In questa faticosa ed incessante ricerca, man mano che i Cavalieri si addentrano in terre sconosciute ed ostili, si annuncia la scoperta di indicatori sempre più vicini all'IPQ, come indizi che guidino sempre più in dettaglio verso il luogo in cui riposa la sacra reliquia. Fra le tappe della ricerca troviamo: il numero di pubblicazioni, le loro citazioni, la VQR individuale, l'Impact Factor, la peer review, e decine di altri indicatori e strumenti eterogenei declinati anche in infinite classificazioni di riviste o ranking di strutture, etc. Per ogni successo reclamato da un Cavaliere e contestato dagli altri, la risposta è sempre la stessa: d'accordo, non è il vero Graal, ma è un passo nella giusta direzione. Finché non troviamo quello vero, usiamo quello che abbiamo: meglio un approssimato (o approssimativo) Graal che non averne nessuno! Come si può vivere senza la certezza che il giusto sarà distinto dal peccatore, il grano dal loglio, l'inattivo dall'eccellente?

Il consenso raccolto dalla necessità di valutare la ricerca è talmente plebiscitario, almeno a parole, da meritare un po' di sano scetticismo illuminista. Vediamo che indizi possiamo elencare sulla possibilità che la valutazione della ricerca sia, o rischi di divenire, una leggenda priva di fondamento.

La ricerca scientifica si occupa di produrre innovazioni che, per definizione, sono diverse dall'esistente, e quindi necessariamente eterogenee ed incomparabili. Al contrario un qualsiasi criterio valutativo necessita di una unità di misura, quindi richiede omogeneità: qualsiasi indicatore che omogenizzi entità eterogenee sarà certamente distorto e non potrà mai rappresentare l'irriducibile unicità di un ricercatore (o prodotto scientifico).

Lo scopo della valutazione è quello di contribuire alla gestione del sistema della ricerca, ma non può sostituirsi ad esso perché è rivolta esclusivamente al passato. Se le politiche strategiche della ricerca si basano solo, o principalmente, sulla valutazione necessariamente si penalizzano le direzioni più innovative, con maggiore potenziale. □ come guidare guardando solo lo specchietto retrovisore: finché la strada è dritta nessun problema, o quasi, ma alla prima curva...



Le procedure di valutazione sono uno strumento di potere, e di conseguenza sono gestite dalle persone con maggior prestigio, che cioè hanno accumulato maggiori successi e con maggiore esperienza e, probabilmente, anzianità. Ma esattamente per questo stesso motivo potrebbero essere proprio i soggetti più ostili verso risultati, o anche solo progetti, che potrebbero minare i risultati che sono alla base della loro posizione. I decisori stessi, quindi, rischiano di avere incentivi opposti allo scopo ultimo della valutazione perché tenderanno a premiare chi conferma i successi passati e non chi prova (minaccia) di superarli. Questo rischio attiene sia le comunità scientifiche che realizzano la (auto)valutazione, con il rischio loro specifico della autoreferenzialità, sia un valutatore esterno che comunque si muove in un contesto di ricerca consolidata.

Questi argomenti, e potenzialmente molti altri, vanno però soppesati contro altri a favore della valutazione, numerosissimi e, in alcuni casi, giustificati. Per quello che interessa qui può però bastare un argomento semplice e molto robusto. Non la chiamiamo valutazione, se non ci piace, ma in regime di risorse finite (per non dire infinite-sime), è necessario fare scelte su quali potenziali progetti vanno finanziati e quali no, chi assumere e chi no, etc. Sia "valutazione" il termine che usiamo, o qualsiasi altro, in ogni caso il semplice fatto che delle decisioni vadano prese richiede comunque un criterio selettivo. Possiamo rinunciare alla fede in un giudizio ultraterreno, ma delle regole di convivenza sociale vanno comunque stabilite e poi rispettate.

Per immaginare l'esistenza senza il conforto di un Santo Graal è necessario partire dal convincimento

che un criterio unico ed universale di qualità della ricerca è semplicemente irrealizzabile. Questo non implica che non sia possibile valutare mai, ma che bisogna rinunciare ad una classifica unica ed assoluta, e concentrarsi sugli obiettivi concreti, che sono differenziati e richiedono criteri di valutazione differenziati, anche radicalmente. Nel mondo accademico si prendono decisioni riguardo la distribuzione di fondi di ricerca per spese correnti, per progetti pluriennali o per posizioni permanenti, assimilabili a spese in conto capitale. A volte interessa la quantità di lavori prodotti, altre volte l'eccellenza. Si può decidere di incentivare la ricerca alla frontiera, estremamente rischiosa ma anche estremamente remunerativa in caso di successo, oppure spingere per avere una produzione di ricerca meno ambiziosa ma dagli esiti più probabili e ed affidabili. Nessuno di questi obiettivi è, di per sé, giusto o sbagliato, ma la scelta di uno o dell'altro riguardano le politiche da adottare. Quello che qui interessa è che questi obiettivi richiedono indicatori diversi, tali che lo stesso soggetto può essere valutato in modo specifico a seconda dell'obiettivo rispetto al quale si svolge la valutazione. In conclusione, la valutazione non è un concetto assoluto, ma può acquisire un significato coerente solo se si specificano gli obiettivi per la quale si svolge, altrimenti si ricade nella fede del potere taumaturgico del Santo Graal che, in forza dei suoi poteri divini, sarebbe in grado di valutare perfettamente le qualità di ciascuno.

Le considerazioni fatte finora non costituiscono un tentativo di troncane il dibattito su indicatori e criteri adatti alla valutazione. Anzi, al contrario, lo vorrebbero alimentare: continuiamo la ricerca del Graal, ma con la coscienza che è la ricerca stessa che interessa, non il suo obiettivo. Quello che serve sono tanti diversi criteri che permettano, conoscendo l'obiettivo, di valutare il (criterio) valutatore. Dato un certo obiettivo è infatti possibile (almeno in linea di massima) verificare quali risultati produrrebbero diversi indicatori proposti e stimare il grado di soddisfazione dell'obiettivo. Le divergenze, ovviamente, non finirebbero mai completamente. Si può discutere su quale obiettivo sia di maggiore rilievo oppure, nel caso di concordanza sull'obiettivo, sui risultati prodotti (o stimati) da ciascun criterio. In fondo, dietro ad ogni processo di valutazione, vi sono anche scelte precise di politica della valutazione. Essere consapevoli di questo già sarebbe un enorme passo avanti, in quanto (almeno) si converrebbe su che cosa, esattamente, si è o non si è d'accordo.



Data la varietà degli obiettivi e dei possibili indicatori, è però necessario essere in grado di applicarli. Questo si può fare solo avendo a disposizione una base di dati "neutri", puramente descrittivi e privi di qualsiasi connotazione ordinale, anche solo implicitamente. Questi dati dovrebbero essere il più completi possibile e di accesso universale, con modalità che permettano facilmente il loro uso per testare diversi criteri e valutarne il risultato. Tale base di dati deve chiaramente indicare tutte le caratteristiche del prodotto in modo che qualsiasi osservatore sia in grado di apprezzarne la natura generale come elemento di una classe (ad esempio, "articolo su rivista con referee"), ma anche permettere l'accesso anche agli aspetti unici e specifici disponibili per ogni prodotto (l'intero set di



metadati il sommario, le citazioni ricevute, etc.). L'infrastruttura del database dovrà prevedere tutti i casi più frequenti di tipi di prodotti scientifici, con controlli (si spera automatizzabili) sulla veridicità del dato inserito. Questa base di dati, pubblica e costantemente aggiornata, deve essere considerata come l'unica fonte completa di dati della ricerca. Il che vuol dire che il profilo di un ricercatore viene definito unicamente e compiutamente dai suoi prodotti (eventualmente integrati da altro materiale, come il cv o ogni altro elemento che il ricercatore stesso reputi utile inserire). I dati di ogni ricercatore sono quindi disponibili nei tempi e nei modi definiti dalle regole di aggiornamento della base di dati, e la sua pubblicità garantisce la completezza, correttezza e rapida correzione di eventuali errori.

Avendo a disposizione una base di dati di tal genere, in sostanza una anagrafe dei prodotti della ricerca e dei ricercatori che li hanno realizzati, sarebbe possibile calcolare qualsiasi indicatore specifico per qualunque obiettivo, rendendo esplicita la differenza tra il profilo completo di un ricercatore ed il valore che lo stesso profilo produce applicando un algoritmo progettato come criterio di valutazione per uno specifico scopo. La disponibilità dei dati e la pubblicazione dell'algoritmo renderebbero possibile a chiunque verificare le conseguenze del criterio usato, opporre obiezioni o proporre alternative, osservando sempre l'effetto generale che ne potrebbe derivare. Similmente, sarebbe possibile facilmente valutare eventuali eccezioni risalendo direttamente alla fonte del dato ipotizzato come anomalo e verificare sia l'esistenza di possibili errori che l'applicabilità di una eccezione, che può essere adeguatamente motivata e giudicata da chiunque.

La disponibilità dei dati renderebbe inutile la necessità di indicatori unici validi per tutti i settori scientifici e per ogni scopo, permettendo a chiunque di sperimentare criteri adatti a diversi contesti, come con un laboratorio del "piccolo Anvur". A titolo di esempio, consideriamo il caso di una commissione di abilitazione. Avendo accesso diretto alla massa dei dati per tutti i candidati, la commissione potrebbe costruirsi criteri decisionali complessi pur mantenendo totale trasparenza sulle motivazioni delle scelte. Ad esempio, scegliere di fissare criteri generali molto selettivi per concedere automaticamente un giudizio positivo ai candidati chiaramente più meritevoli, e poi studiare i risultati prodotti da diversi criteri specifici per individuare profili di candidati particolari, meno appariscenti ma egualmente meritevoli di giudizio positivo. Questo metodo darebbe la possibilità di evitare indicatori grezzi ed universali, e quindi sicuramente inadatti per alcuni settori e profili. Al contrario, un concorso comparativo o una commissione di premialità, con un numero limitato di candidati da considerare, si potrebbe permettere di confrontare i profili integrali dei candidati senza dover ricorrere a indicatori sintetici concepiti per altro scopo. Che il sistema adottato sia la peer review o uno stimatore bibliometrico, la pubblicità dei dati originali è il solo mezzo con cui limitare ex-ante le inevitabili controversie, che vengono invece oggi sanate ex-post in una molteplicità di modi.



Come piccola digressione, ricordiamo che l'intrinseca imprevedibilità dell'impatto futuro della ricerca richiede che l'obiettivo di un sistema valutativo debba perseguire non solo la premialità delle ricerche che appaiono al momento più promettenti, ma anche il mantenimento della varietà di punti di vista, metodi ed approcci che, nel lungo periodo, è necessaria per affrontare l'incertezza delle esigenze future della società. Come nel caso delle scelte in agricoltura, la monocoltura della varietà più produttiva al momento fornisce sicuramente maggiori ritorni immediati, ma condanna nel lungo periodo il sistema a perdere la sua capacità di adattamento. Il miglior esempio di questa necessità è lo sviluppo dell'intero settore dell'information technology che è stato reso possibile solo grazie al lavoro di un logico morto nel 1864...

Concludendo, allo scopo di basare le discussioni sulla valutazione della qualità della ricerca su fondamenta robuste, e, se non proprio oggettive, almeno largamente condivise, è fortemente auspicabile implementare nel più rapido tempo possibile e nel modo migliore in termini di efficacia ed efficienza, l'anagrafe della produzione scientifica secondo linee definite in termini puramente tecnico-organizzativi, prive di qualsiasi risvolto valutativo diretto. Tale anagrafe deve solo assicurare di contenere tutte le informazioni possibili sulla produzione scientifica dei ricercatori registrati senza alcun risvolto valutativo. Ad esempio, dovrà indicare l'articolo sulla rivista XYZ, non la classe di qualità della rivista. Chi ne sentisse la necessità potrebbe poi interrogare l'anagrafe su tutti gli articoli della

rivista contenuti nell'anagrafe, o i nomi dei ricercatori che vi hanno pubblicato. L'anagrafe è la fonte dei dati, selezionati esclusivamente sul fatto di essere reclamati da un ricercatore come prodotto del suo lavoro. Le valutazioni, diverse in funzione di scopi diversi, saranno costruite come elaborazione di questi dati.

Non è qui il caso di entrare nei dettagli tecnici e dei costi necessari per l'implementazione di una anagrafe della ricerca, ma vale la pena ricordare che esistono, in ambito internazionale, una serie di standard rispetto agli identificativi univoci (almeno di persone e pubblicazioni). Inoltre in ambito nazionale è stata realizzata una consultazione pubblica su iniziativa del CUN ed in collaborazione con l'Anvur volta ad identificare i criteri di scientificità delle pubblicazioni, e da questa consultazione sono seguite le linee guida elaborate dal CUN che dovrebbero fornire la base per Anpreps. Purtroppo, il processo per giungere, in modo condiviso nella professione, alla realizzazione dell'Anagrafe nazionale nominativa dei professori e dei ricercatori e delle pubblicazioni scientifiche, prevista dalla normativa sin dal gennaio 2009, sembra essersi arenato, anche se tutti i soggetti potenzialmente coinvolti nella sua realizzazione continuano a sostenerne la necessità. Vogliamo quindi riaffermare l'urgenza improcrastinabile di poter disporre di tale strumento come pre-condizione necessaria a qualsiasi discussione riguardo metodi di valutazione e suoi risultati.



Copyright :

All this contents are published under [Creative Commons Attribution-NonCommercial-ShareAlike 2.5 Generic License](https://creativecommons.org/licenses/by-nc-sa/2.5/).

for reproduced, please specify from this website [ROARS](http://www.roars.it/online/) AND give the URL.

Article link : <http://www.roars.it/online/?p=37294>