



UNIVERSITÀ DEGLI STUDI DI TORINO

Nucleo di Valutazione

Contributo al dibattito sulla valutazione dei prodotti della ricerca

Il Nucleo di Valutazione esprime il proprio apprezzamento per il progressivo diffondersi della cultura della valutazione all'interno dell'Ateneo e per l'attribuzione delle risorse su base valutativa, ma le prime applicazioni operative richiedono un momento di riflessione che consenta di proseguire sulla strada intrapresa intervenendo, dove necessario, con le opportune, e probabilmente necessarie, correzioni.

Il presente documento si sofferma pertanto su alcuni elementi di criticità riscontrati dal Nucleo di Valutazione al termine di questo primo anno di lavoro.

Analisi della situazione attuale

La tabella dei pesi dei prodotti della ricerca

L'attuale sistema di pesi è stato proposto ed adottato dal Nucleo di valutazione fin dalla Relazione annuale 2001 e, da allora, la tabella ha subito relativamente poche variazioni ed è stata applicata ogni anno per la valutazione della produzione scientifica dei dipartimenti presentata nell'ambito della relazione annuale sulla ricerca.

Nel corso del 2007 il gruppo di lavoro costituito dal Vice Rettore per i sistemi informativi, da alcuni componenti della Commissione Ricerca del Senato Accademico (CRSA) e da un rappresentante del Nucleo, ha concordato di adottare una nuova suddivisione delle tipologie di pubblicazioni scientifiche in vista della realizzazione del Catalogo dei prodotti di Ateneo.

Conseguentemente il Nucleo ha rivisto la propria tabella di valutazione, adattandola alle nuove tipologie di prodotti.

Dal 2009 la medesima tabella è stata utilizzata dalla CRSA per misurare la produzione scientifica dei singoli docenti ai fini dell'allocazione di risorse.

Nella versione attuale della tabella, i prodotti sono classificati in nove "tipologie", ulteriormente suddivise in "sub-tipologie" e valutati da 0,3 ad 3 punti, in modo differenziato tra le macro-aree delle Scienze della Natura e delle Scienze Umane (Tabella 1).

Il Catalogo dei prodotti della ricerca

Fino al 2007, l'elenco dei prodotti della ricerca sul quale si basano le valutazioni è stato trasmesso direttamente dai dipartimenti al Nucleo di Valutazione, utilizzando di volta in volta supporti diversi. Dal 2008 viene utilizzato il database del Catalogo della ricerca che ad oggi raccoglie i prodotti di ricerca inseriti personalmente da docenti di ruolo (PO, PA e RU) e assegnisti; è in corso di attuazione l'inserimento dei prodotti anche da parte dei dottorandi. L'ultima estrazione dei prodotti del catalogo, avvenuta in data 04/06/2010, ha evidenziato la presenza, limitatamente al quinquennio 2005-09 e al netto delle duplicazioni, di 35.721 prodotti, articolati per le aree delle Scienze della Natura e delle Scienze Umane come illustrato nella Tabella 2.



UNIVERSITÀ DEGLI STUDI DI TORINO

Nucleo di Valutazione

I sistemi di valutazione

Parallelamente, ma concettualmente distinti dall'implementazione del Catalogo, sono stati sviluppati dal Nucleo di Valutazione e dalla Commissione Ricerca Scientifica del Senato Accademico due sistemi di valutazione che, pur condividendo la medesima classificazione e il medesimo meccanismo di ponderazione dei prodotti, si differenziano per obiettivi e metodologia.

Il sistema di valutazione del Nucleo di Valutazione ha come oggetto prioritario la valutazione delle strutture. I prodotti vengono quindi aggregati per dipartimento, eliminando i doppioni "intra-struttura". La valutazione di ciascuna struttura è determinata dalla somma dei prodotti afferenti alla struttura medesima, opportunamente ponderati in base ai coefficienti di ciascuna tipologia di prodotto.

Il sistema adottato dalla Commissione Ricerca del Senato Accademico ha invece come oggetto prioritario di valutazione i singoli docenti e, solo indirettamente, le loro strutture di afferenza (dipartimento, aree scientifico-disciplinari o dottorati). A ciascun docente viene attribuito un punteggio che corrisponde alla somma dei punti dei prodotti di cui è autore o co-autore, nel quinquennio considerato per la valutazione. Per evitare che la somma di prodotti di minore rilevanza scientifica diventi predominante è stato posto un tetto al punteggio raggiungibile con le "altre pubblicazioni" (2,7 punti).

I docenti del 1° quartile di ciascuna delle 16 aree dell'Ateneo sono collocati in fascia A (25% dei docenti di ciascuna area), quelli del 2° e 3° quartile sono collocati in fascia B (50% dei docenti di ciascuna area) e infine quelli del 4° quartile sono collocati in fascia C purché il loro punteggio sia uguale o superiore a 3.

I docenti il cui punteggio nel quinquennio è inferiore a 3 punti sono considerati scientificamente non attivi e collocati in fascia D.

Criticità, osservazioni e proposte sui sistemi attuali di valutazione

La distribuzione dei docenti in fasce

Occorre sottolineare che il modello di valutazione adottato dalla CRSA non prevede di fare distinzioni inter-area se non dovute al numero di docenti collocati in fascia D.

La dotazione alle aree è quindi rimasta invariata al netto dei docenti in fascia D, mentre la dotazione ai dipartimenti non è più dipesa solo del numero degli afferenti ma dalla loro distribuzione nelle fasce A, B e C. Per quanto riguarda i dottorati, si è considerata la proporzione di docenti delle fasce senza distinzione tra A e B.

Docenti in fascia D

I docenti in fascia D rappresentano il 10% del totale e sono distribuiti pressoché ugualmente tra area scientifica ed umanistica. Alcuni docenti non hanno inserito alcun prodotto, altri non hanno chiuso tutti loro prodotti, tuttavia, nella maggior parte dei casi, il Nucleo ritiene che il sistema attuale abbia con buona approssimazione identificato docenti scientificamente non attivi.

Docenti in fascia A, B, C

Per ciò che riguarda la collocazione dei docenti nella fasce A, B e C, essa trova un certo riscontro nell'analisi condotta sui prodotti di maggior peso. Tuttavia si possono verificare delle distorsioni,



UNIVERSITÀ DEGLI STUDI DI TORINO

Nucleo di Valutazione

perché il sistema attuale favorisce produzioni scientifiche quantitativamente elevate a scapito di produzioni meno numerose ma di maggiore qualità. Questa criticità è riscontrabile sia nell'area scientifica, sia nell'area umanistica ed è ritenuta particolarmente grave da parte del Nucleo.

Per l'area scientifica una motivazione è senz'altro legata alla tripartizione delle riviste ISI con *Impact Factor* che prevede fasce eccessivamente ampie con la sopravvalutazione di numerose riviste. Per ciò che riguarda l'area umanistica, una spiegazione può ricondursi all'assenza di distinzioni nel punteggio tra pubblicazioni che, pur appartenendo a una medesima tipologia, hanno valore molto diverso.

Confrontabilità inter-area e intra-area: eterogeneità dei prodotti e multiple authorship

L'eterogeneità della produzione scientifica rappresenta un'obiettivo difficoltà nell'applicazione di sistemi di valutazione basati su indicatori di natura meramente bibliometrica. Le differenze negli obiettivi di ricerca influenzano, in parte per ragioni strutturali, in parte per tradizioni storiche, alcuni elementi dei meccanismi di produzione della ricerca che a cascata rendono difficile la costruzione di un sistema di valutazione unitario. Questi elementi comprendono, ma non si limitano a: i) tipologia del prodotto (e.g. articolo scientifico versus monografia), ii) area di diffusione del prodotto (e.g. internazionale versus nazionale) e lingua di comunicazione (e.g. inglese versus lingua nazionale), iii) numero di autori per prodotto (e.g. autore singolo versus gruppi di ricerca).

Questa eterogeneità si riscontra tra aree fra loro lontane, ma anche all'interno di una stessa area, qualora questa non coincida con un insieme di docenti che afferiscono alla medesima comunità scientifica e che quindi non condividono la medesima tipologia di produzione scientifica.

L'approccio attualmente in vigore presso l'Ateneo riconosce in modo molto limitato, e insufficiente nell'opinione del Nucleo, questa eterogeneità, limitandosi ad attribuire punteggi diversi alla medesima tipologia di prodotti in base all'appartenenza del docente alla macro-area di scienze umane o alla macro-area di scienze della natura.

Il riconoscimento dell'esistenza di una ampia varietà di tecnologie di produzione pone come problema prioritario quello della comparabilità dei risultati. La via maestra per la valutazione delle aree di ricerca e dei dipartimenti dell'Ateneo passa sicuramente per il confronto con strutture comparabili (*benchmarking*), da condurre ove possibile a livello internazionale o nazionale. A questo fine, i dati che saranno resi disponibili dal CIVR al termine della prossima valutazione quinquennale (2004-08) della ricerca, saranno sicuramente un punto di riferimento importante.

Dal momento, però, che le valutazioni nazionali avvengono con scadenza pluriennale e che risulta al momento attuale utopico che l'Ateneo destini le risorse necessarie per una sistematica attività di *benchmarking*, qualora nei prossimi anni si intendano utilizzare meccanismi di valutazione inter-area basati sul Catalogo dei prodotti sarà essenziale tenere conto sia dell'eterogeneità della produzione scientifica sia del problema specifico della *multiple authorship* per la costruzione di un modello di valutazione equo.

La natura stessa della ricerca nel macrosettore delle scienze della natura richiede generalmente la collaborazione di molti ricercatori mentre il lavoro di ricerca nelle scienze umane è maggiormente condotto a livello individuale o in piccoli gruppi. Motivo per cui la confrontabilità inter-area, e in alcuni casi anche intra-area, richiede come presupposto imprescindibile che il contributo dei prodotti della ricerca alla valutazione del singolo docente venga standardizzato in funzione del numero degli autori del prodotto. Esiste una consolidata tradizione internazionale, in cui il punteggio attribuito a ciascun prodotto viene standardizzato applicando un fattore di correzione non



UNIVERSITÀ DEGLI STUDI DI TORINO

Nucleo di Valutazione

lineare nel numero degli autori. Esistono anche riferimenti nazionali quali ad esempio quelli applicati da alcune aree disciplinari o dalla prossima valutazione quinquennale della ricerca del CIVR.

Il problema della *multiple authorship* è stringente nel caso di valutazioni inter-aree ma esiste anche nelle valutazioni intra-area qualora l'area non coincida con un insieme omogeneo di ricercatori. Nel bilanciare eventuali criteri legati alla *multiple authorship*, occorre però porre attenzione poiché si potrebbe correre, in particolare per alcune aree scientifiche, il rischio di sovrastimare lavori come per esempio le *review* e contemporaneamente sottostimare il valore di contributi quali quelli del primo e ultimo autore di articoli di grande impatto scientifico.

Incentivi e comportamenti opportunistici

I meccanismi di valutazione, soprattutto se utilizzati per l'allocazione delle risorse, modificano il sistema degli incentivi individuali e possono indurre a comportamenti opportunistici invece di promuovere il miglioramento.

È opinione del Nucleo di Valutazione che gli attuali schemi di riferimento inducano, soprattutto nei ricercatori, comportamenti volti a privilegiare la "quantità" rispetto alla "qualità" e che conseguentemente si possa produrre una distorsione verso il basso nella produzione scientifica. La ragione è da ricercarsi in primo luogo nella insufficiente variabilità dei punteggi tra le diverse tipologie di prodotto. Questo vale sia per le aree afferenti alle scienze della natura, in cui tre articoli "bronze" equivalgono ad un articolo sulla migliore rivista dell'area, sia soprattutto per le aree di scienze umane, in cui tre articoli su riviste nazionali senza IF (e magari senza alcun processo di controllo da parte di *referee*) equivalgono ad una monografia pubblicata dalla migliore casa editrice accademica internazionale.

Inoltre l'elevato numero di tipologie di prodotti e la conseguente elevata numerosità dei prodotti da valutare rende impossibile, o comunque eccessivamente costoso, il monitoraggio da parte degli uffici preposti e la conseguente correzione di eventuali errori materiali o di interpretazione. Inoltre, la percepita mancanza di monitoraggio può indurre comportamenti non corretti quali l'allocazione del medesimo prodotto in più di una tipologia o in una tipologia non coerente con la natura del prodotto stesso.

Riduzione delle tipologie e/o del numero di prodotti sottoposti a valutazione

La riduzione dei prodotti da valutare può essere realizzata sia attraverso una riduzione delle tipologie ammissibili, sia attraverso l'introduzione di una soglia massima di prodotti per docente¹.

Per entrambi gli aspetti, il Nucleo suggerisce di utilizzare un modello coerente, se non uguale, a quello adottato per la valutazione quinquennale della ricerca del CIVR.

La riduzione delle categorie e la selezione di un numero limitato di prodotti (e.g. dieci/cinque prodotti su un arco di tempo quinquennale)² da parte di ciascun docente innescherebbe un meccanismo di autoselezione dei prodotti "migliori" e agevolerebbe sia l'attività di *monitoring* da parte degli uffici preposti, sia l'integrazione con altre fonti tipicamente utilizzate per il calcolo del numero delle citazioni quale ad esempio l'indice di Hirsch.

¹ Numero massimo equivalente nel caso in cui si applichi un fattore di co-authorship (in analogia alla valutazione quinquennale della ricerca del CIVR)

² *Idem*



UNIVERSITÀ DEGLI STUDI DI TORINO

Nucleo di Valutazione

Verifiche dell'attendibilità del Catalogo dei prodotti della ricerca

Attualmente, la presenza di errori nel Catalogo della ricerca ha contribuito a creare distorsioni nella valutazione e in qualche caso ha generato anomalie particolarmente evidenti nei punteggi.

Il Nucleo ritiene indispensabile definire con maggiore chiarezza le tipologie e le loro caratteristiche per evitare fraintendimenti o errate classificazioni dei prodotti, ma auspica anche che cresca il senso di responsabilità individuale nell'inserimento dei dati nel Catalogo della ricerca e che siano adottate delle politiche in grado di garantire l'attendibilità del database, presupposto fondante di ogni forma di valutazione.

La trasparenza del processo valutativo

Il Nucleo ha ritenuto prioritario promuovere lo sviluppo e l'implementazione di una piattaforma informatica per la diffusione dei risultati della valutazione (www.ndv.unito.it). La piattaforma è stata messa a disposizione dei Direttori di Dipartimento e dei Rappresentanti di Area all'inizio del 2010, sia per assicurare la trasparenza delle valutazioni che per promuovere la segnalazione di errori facilmente identificabili. La piattaforma è stata in seguito ulteriormente sviluppata per permettere a ciascun docente di vedere l'esito delle proprie valutazioni.

Il Nucleo auspica che l'Ateneo prosegua lungo la strada tracciata ritenendo la trasparenza un elemento importante del processo di valutazione.

Rispetto dei vincoli di costi e benefici

Come accennato in diverse parti del presente documento, l'elaborazione di nuovi criteri per la valutazione della produzione scientifica deve tenere conto della sostenibilità delle procedure valutative rispetto alle risorse finanziarie da distribuire e alle risorse umane che sono impiegabili in queste attività.

Il Nucleo ritiene che i meccanismi su cui si basa la valutazione debbano continuare ad essere basati su meccanismi bibliometrici che consentano un ampio utilizzo dei mezzi informatici, almeno fin tanto che non sia operativo un vero osservatorio della ricerca e che i prodotti sottoposti a valutazione siano certificati dagli autori, sempre in analogia alla valutazione quinquennale della ricerca del CIVR.

La revisione dei meccanismi di ponderazione

Per migliorare il processo selettivo e incentivare la ricerca di frontiera il Nucleo ritiene opportuno procedere ad una maggiore differenziazione del peso attribuito alle tipologie e sub-tipologie dei prodotti. Questa operazione può essere effettuata abbastanza agevolmente nel caso di riviste dotate di *impact factor*. Risulta invece di più difficile attuazione nel caso di libri, capitoli su libri e riviste senza *impact factor*. In linea di principio, infatti, l'attribuzione di pesi diversi richiede una classificazione ordinale, almeno in categorie, di case editrici e di riviste.

La predisposizione di schemi di classificazione per area costituisce, a parere del Nucleo, un presupposto necessario per modificare il sistema di valutazione nelle aree in cui la pubblicazione su riviste con *impact factor* non rappresenta il veicolo più comune di diffusione dei risultati del lavoro scientifico.

Per alcune aree disciplinari esistono classificazioni pubblicamente disponibili realizzate a livello internazionale o nazionale, talvolta dalle società scientifiche di riferimento. Queste tuttavia differiscono per ampiezza, obiettivi e, in alcuni casi, impostazione culturale. Costituiscono dunque informazioni utili ma non esaustive.



UNIVERSITÀ DEGLI STUDI DI TORINO

Nucleo di Valutazione

Nel 2008 il Nucleo aveva proceduto ad una prima raccolta di informazioni relative alla classificazione delle riviste e case editoriali presso i dipartimenti. Tali informazioni sono in fase di elaborazione e potranno costituire una base di partenza.

Infine potranno essere presi in esame anche i parametri che verranno definiti a breve dai diversi *panel* previsti dalla prossima valutazione quinquennale della ricerca.

L'incentivazione al miglioramento

Per quanto riguarda la necessità di incentivare il miglioramento nel rispetto dell'impostazione attuale, il Nucleo osserva che si potrebbero, al termine del processo di revisione del presente sistema di valutazione, declinare le fasce in livelli (ad esempio A1, A2, A3, B1...) in modo tale da evidenziare indici, anche parziali, di miglioramento.

Sarebbe inoltre auspicabile calcolare i quartili di ciascuna area in base al numero di docenti attivi e non in base al numero totale di docenti, evitando così di avere aree nelle quali la categoria C è praticamente inesistente in quanto circa il 25% dei docenti risulta in fascia D.

Definizione degli obiettivi

La valutazione non è esente da implicazioni politiche. Sistemi di valutazione diversi (identificazione di docenti scientificamente non attivi, produzione media delle aree, produzione di eccellenza...) rispondono a obiettivi diversi. Il Nucleo sollecita quindi la definizione degli obiettivi da parte degli organi di governo perché si possa procedere alla definizione di sistemi di valutazione coerenti.

Conclusioni

La revisione del sistema di valutazione richiede la condivisione di obiettivi e strategie all'interno dell'Ateneo. Le fasi operative costituiscono un processo complesso e molto oneroso in termine di tempo che difficilmente possono essere affidate al Nucleo tenuto conto delle sue sempre crescenti incombenze istituzionali. Il Nucleo auspica dunque l'attuazione di un gruppo di lavoro dedicato e rappresentativo di tutte le aree che possa in tempo ragionevole portare a compimento la definizione e l'attuazione di un sistema di valutazione che promuova il miglioramento nel rispetto della pluralità.

Il Nucleo di Valutazione dell'Università di Torino, a.a. 2009/2010:

Prof.ssa Isabelle PERROTEAU, Prof. Fausto GORIA, Dott.ssa Ines FABBRO, Prof.ssa Silvia GIORDANO, Prof. Andrea SCHUBERT, Prof. Alessandro SEMBENELLI, Prof. Francesco TUCCARI



UNIVERSITÀ DEGLI STUDI DI TORINO

Nucleo di Valutazione

Tabella 1

TIPOLOGIE CATALOGO		Sc. Natura	Sc. Umane			
LIBRO						
1a	Monografia o trattato scientifico	1	3			
1b	Concordanze, indici, bibliografia	0,3	0,3			
1c	Edizione Critica, pubblicazione di fonti inedite	1	1			
CONTRIBUTO SPECIFICO IN VOLUME						
2a	Capitolo di libro, Articolo in volume miscellaneo	1	1			
2b	Voce di Enciclopedia, Dizionario, Atlante o Carta tematica	0,3	0,3			
2c	Prefazione, Postfazione o Introduzione	0,3	0,3			
PUBBLICAZIONE IN RIVISTA		Senza IF	Senza IF	IF³ gold	IF⁴ silver	IF⁵ bronze
3a	Articolo su rivista	0,3	1	3	2	1
3b	Nota a sentenza firmata	0,3	0,3			
3c	Recensione, scheda o segnalazione bibliografica, Replica, Commento o simili	0,3	0,3			
PROCEEDINGS						
4a	Proceedings (in volume)	0,3	0,3			
4b	Proceedings (in rivista)	0,3	0,3			
CURATELA						
5a	Cura di volume collettaneo o di atti di convegno	0,3	0,3			
5b	Cura di cataloghi	0,3	0,3			
5c	Cura di numeri monografici di rivista	0,3	0,3			
TRADUZIONE						
6a	Traduzione di libro	0,3	0,3			
6b	Traduzione di capitolo di libro o contributo in atti di convegno (in volume)	0,3	0,3			
6c	Traduzione di pubblicazione (in rivista) o contributo in atti di convegno (in rivista)	0,3	0,3			
SOFTWARE E PRODOTTI MULTIMEDIALI						
7a	Prodotto Software open source	0,5	0,5			
7b	Prodotto multimediale	0,5	0,5			
MANUFATTO O OPERA D'ARTE						
8a	Prodotto Ing. Civile e Architettura	0,3	0,3			
8b	Prodotto artistico o spettacolare (documentato da rassegna stampa)	0,3	0,3			
8c	Installazione audiovisiva e/o interattiva (documentata da rassegna stampa)	0,3	0,3			
9	<u>Brevetto</u>	1	1			
10	<u>Altro</u>	0	0			

³ Per "IFgold" si intende articolo pubblicato su riviste appartenenti al 1° quartile della categorie di appartenenza

⁴ Per "IFsilver" si intende articolo pubblicato su riviste appartenenti al 2° o 3° quartile della categorie di appartenenza

⁵ Per "IFbronze" si intende articolo pubblicato su riviste appartenenti al 4° quartile della categorie di appartenenza



UNIVERSITÀ DEGLI STUDI DI TORINO

Nucleo di Valutazione

Tabella 2

TIPOLOGIE CATALOGO		Sc. Natura	Sc. Umane
<u>LIBRO</u>			
1a	Monografia o trattato scientifico	201	912
1b	Concordanze, indici, bibliografia	-	12
1c	Edizione Critica, pubblicazione di fonti inedite	2	66
<u>CONTRIBUTO SPECIFICO IN VOLUME</u>			
2a	Capitolo di libro, Articolo in volume miscelaneo	1.197	4.817
2b	Voce di Enciclopedia, Dizionario, Atlante o Carta tematica	33	798
2c	Prefazione, Postfazione o Introduzione	44	509
<u>PUBBLICAZIONE IN RIVISTA</u>			
3a	Articolo su rivista	10.441	3.354
	<i>di cui su riviste con IF Gold</i>	4.384	111
	<i>di cui su riviste con IF Silver</i>	2.915	169
	<i>di cui su riviste con IF Bronze</i>	613	49
	<i>di cui su riviste senza IF</i>	2.529	3.025
3b	Nota a sentenza firmata	-	635
3c	Recensione, scheda o segnalazione bibliografica, Replica, Commento o simili	327	1.393
<u>PROCEEDINGS</u>			
4a	Proceedings (in volume)	4.613	961
4b	Proceedings (in rivista)	1.530	79
<u>CURATELA</u>			
5a	Cura di volume collettaneo o di atti di convegno	244	705
5b	Cura di cataloghi	-	27
5c	Cura di numeri monografici di rivista	27	90
<u>TRADUZIONE</u>			
6a	Traduzione di libro	8	80
6b	Traduzione di capitolo di libro o contributo in atti di convegno (in volume)	22	133
6c	Traduzione di pubblicazione (in rivista) o contributo in atti di convegno (in rivista)	-	46
<u>SOFTWARE E PRODOTTI MULTIMEDIALI</u>			
7a	Prodotto Software open source	41	12
7b	Prodotto multimediale	63	59
<u>MANUFATTO O OPERA D'ARTE</u>			
8a	Prodotto Ing. Civile e Architettura	-	-
8b	Prodotto artistico o spettacolare (documentato da rassegna stampa)	-	16
8c	Installazione audiovisiva e/o interattiva (documentata da rassegna stampa)	8	4
9	<u>Brevetto</u>	113	1
10	<u>Altro</u>	1.241	955

N.B. 98 prodotti sono stati fatti da coautori appartenenti ad entrambe le macro-aree e quindi sono stati duplicati nella tabella proposta