Appunti per la presentazione del 21 marzo 2016

1) Sorvegliare e Punire?

Quando chi definisce i criteri, chi esercita la funzione di controllo e chi eroga le sanzioni è un unico soggetto viene a cadere il meccanismo di *checks and balances* che sta alla base di qualunque sistema anche *lato sensu* democratico.
In questo caso la valutazione si traduce, anche al di là delle intenzioni soggettive dei valutatori, in un esercizio di potere "assoluto" e privo di controllo che, la Storia ci insegna, finisce per impedire il pieno dispiegamento delle potenzialità di un sistema o di un gruppo sociale.
Corollario della logica "punitiva" della valutazione è il *Matthew effect* ("A chi ha sarà dato e a chi non ha sarà tolto") che nel contesto universitario si traduce in un aumento degli squilibri territoriali e in una oggettiva riduzione dell'*humus* scientifico diffuso dal quale soltanto possono nascere anche le "eccellenze".
Pensare che si possano ottenere risultati di altissima qualità concentrando le risorse in poche sedi privilegiate e abbandonando a se stesso il esto del sistema significa non avere alcuna consapevolezza dei meccanismi e delle dinamiche reali che fanno emergere le vocazioni e le competenze all'interno del corpo sociale. Occorre molta ricerca per avere buona ricerca.

2) Misurare e Orientare?

Non ci sono dubbi sul fatto che il sistema dell'alta formazione e della ricerca ha tuttavia bisogno di essere valutato, e che non sia inappropriato cercare di individuare indicatori che offrano una misura dell'efficienza delle strategie e delle soluzioni adottate dalle diverse realtà operanti all'interno del sistema.
Le risorse per definizione non sono infinite, e un orientamento verso il miglior uso delle limitate risorse disponibili è quindi certamente benvenuto.
Sarebbe bene tuttavia non dimenticare che esistono già meccanismi "interni" di valutazione e di autovalutazione.
In particolare la ricerca è continuamente esposta a processi di confronto, in molti casi internazionale e comunque sempre almeno nazionale, che si articolano nella *peer review* dei lavori sottoposti per la pubblicazione, nella selezione degli inviti ai congressi, nelle procedure concorsuali di ogni ordine e grado. Il principale compito di un valutatore "terzo" dovrebbe essere quello della verifica della coerenza degli esiti dell'autovalutazione (che si traduce poi di fatto nelle strategie operative adottate) con quelli di una valutazione "dei pari" già largamente disponibile. "Inventare" nuovi indicatori è in parte segno di una forma di arroganza intellettuale, ed espone anche al rischio di effettuare "esperimenti sociali" *in vivo* senza averne prima verificato *in vitro* i limiti e la possibile "tossicità"

3) I Limiti della Misurazione

In primo luogo non si deve mai dimenticare che, come ha insegnato la fisica microscopica, la misura perturba sempre il sistema, e quanto maggiore è la pretesa di precisione, tanto maggiore è la perturbazione. Negli organismi viventi esistono soglie sopra le quali l'invasività del processo di misura distrugge il fenomeno stesso che si vuole misurare.
In questo contesto non appare sensata la pretesa di arrivare a un *ranking* delle strutture valutate (ovvero a un ordinamento totale e granulare), mentre invece ci si dovrebbe limitare a un *rating* (ovvero all'individuazione di una classe di qualità, all'interno della quale non si possono fare graduatorie, limitandosi al più a 5-6 classi).

Bisogna poi tener presente quello che in teoria dell'errore è definito "errore statistico", per cui quando si analizza un numero limitato di dati l'imprecisione dello strumento di valutazione si traduce in una "varianza" dei risultati non imputabile a differenze reali.
Questa considerazione risulta fondamentale quando si faccia riferimento all'uso di indicatori "oggettivi" in relazione alle valutazioni individuali, per le quali l'ampiezza della fluttuazione statistica è sostanzialmente pari alla media dei valori misurati. In altri termini senza una "lettura critica" del dato (che solo un revisore esperto può fare) l'indicatore potrebbe essere totalmente privo di significato reale.
Viceversa quando si analizza un grande numero di dati la varianza si riduce, ma a quel punto il risultato tende a convergere quasi inevitabilmente a un valore medio, con un meccanismo che solo raramente da luogo a osservabili differenze qualitative.
Infine non bisogna dimenticare che lo strumento di misurazione è 'costruito" da un osservatore che nella definizione stessa dello strumento rischia di introdurre un pre-giudizio legato alle proprie aspettative sull'esito della misura (errore sistematico): gli esempi si potrebbero moltiplicare (ma per capirci basti pensare al sistema tolemaico)

4) I Limiti dell'Orientamento

Una valutazione che sia troppo fortemente condizionata dal riscontro quantitativo dell'impatto (misurato nel caso della bibliometria da IF e conteggi di citazioni, comunque combinati tra loro) spinge inevitabilmente la ricerca verso quegli ambiti nei quali, talvolta per moda contingente, tale riscontro è più facilmente conseguibile. Anche una differenziazione sulla base della *Subject Category* (SC) non è i grado di attenuare il fenomeno, perché se da un lato rende confrontabili anche settori profondamente diversi (*cfr* Radicchi *et al.*) dall'altro non coglie la varietà di situazioni, di ambiti e di profili presenti all'interno di una stessa SC.
I giovani ne risultano pertanto dis-orientati, in quanto le loro scelte saranno sempre più condizionate non tanto dalle reali vocazioni, attitudini e interessi quanto dalle opportunità offerte da un mercato del lavoro nella ricerca nel quale il successo valutativo è misurato da un consenso sostanzialmente quantitativo.
Inevitabile quindi assistere a comportamenti opportunistici e adattivi, non soltanto da parte dei giovani ma anche da chiunque abbia aspirazioni di carriera o anche soltanto di prestigio accademico.
Inevitabile anche un'ulteriore restrizione degli spazi di autonomia non soltanto degli individui ma anche delle istituzioni, che saranno spinta a sostenere con molta più forza le iniziative capaci di produrre un ritorno valutativo del tipo sopra indicato rispetto a proposte anche ambiziose, stimolanti e originali ma incapaci di attrarre attenzione e consensi nel breve periodo.
L'esito potrebbe essere la progressiva scomparsa degli "ecotoni", ovvero quelle zone di transizione nelle quali la grande varietà "biologica" offre da un lato maggiori *chance* di sopravvivenza e dall'altro maggiore spazio per significativi processi evolutivi.

5) Per una nuova primavera

Il primo obiettivo dovrebbe essere quello di una valutazione non "invasiva" (nel senso con cui il vocabolo è utilizzato in medicina) e quindi per quanto possibile non distruttiva del tessuto culturale e sociale esistente e della capacità delle comunità accademiche di costituire "ecotoni" come indicato in precedenza.
Tale obiettivo passa anche attraverso percorsi di condivisione dei criteri di valutazione, che non possono essere applicati ai soggetti e ai sistemi valutati senza che tali soggetti e sistemi abbiano potuto esplicitare le logiche che hanno portato alla definizione del proprio *modus*

operandi. La terzietà dei valutatori, certamente sacrosanta, non deve tradursi in una forma di estraneità alle premesse culturali che stanno alla base dei modi con cui le comunità scientifiche si auto-organizzano, si confrontano al proprio interno e si auto-giudicano. Normalmente il consenso, informale ma diffuso, sui valori e sulle "gerarchie" qualitative tra singoli e tra istituzioni è spesso più accurato di quello che si può estrarre da indicatori che costituiscono comunque sempre una *proxy* quantitativa della qualità reale (e per definizione non "misurabile", ma al più "ordinabile" mediante una scala di giudizi).
In ogni caso, quando la necessità di giungere a un qualche tipo di parametrazione dei giudizi non fosse eludibile, si dovrebbe per lo meno avere il massimo rispetto per le leggi della statistica, tenendo adeguatamente conto di tutte quelle cause di errore che abbiamo già discusso in precedenza e presentando i risultati in una forma che tenga esplicitamente conto di questo margine di errore e della relativa necessaria "tolleranza" (nel senso ingegneristico del termine).
Da ultimo sarebbe importante imparare a far uso dei nuovi strumenti di analisi che si vanno rendendo progressivamente disponibili, in particolare con l'avvento delle indagini basate sui "*big data*" (che non significa "molti dati"). Nel caso che ci interessa, premessa fondamentale sarebbe la costruzione di adeguate, e autonome, basi di dati relativi al sistema universitario (prime fra tutte l'Anagrafe delle Ricerche – ANPRePS – e l'Anagrafe degli Studenti).
Su basi di dati sufficientemente ampie si potrebbe procedere a ricognizioni "a statistica controllata" e quindi con plausibili stime di errore.
Ancor più ambizioso e futuribile, ma non fuori dalla portata delle indagini da volgersi nei prossimi anni, sarebbe un approccio basato sulla *sentiment analysis*, in cui l'enorme insieme non strutturato delle informazioni accessibili in Rete potrebbe offrire indicazioni, in ultima analisi più attendibili di quelle ricavate con i metodi attuali, sul reale impatto sul complesso della società (o su suoi specifici segmenti) del lavoro svolto nelle istituzioni accademiche (sia per la didattica che per la ricerca e la terza missione).