

Sempre di più si richiede un'analisi della produzione scientifica di nazioni, università, istituti di cultura e ricerca, e singoli individui. I ricercatori provetti, che siano agli inizi od in mezzo alla loro carriera, debbono sempre più confrontarsi con questa analisi in termini quantitativi, che inevitabilmente si basa su dati bibliometrici, cioè le pubblicazioni e le citazioni che esse attirano.

Questo *manualetto* si prefigge di dare delle indicazioni di dove e come trovare i parametri di riferimento per quantificare il calibro scientifico, e quindi 'prestigio', delle nazioni ed istituti in cui potreste andare a lavorare, e di voi stessi ricercatori - per definire il vostro profilo da mettere nel CV, e magari discutere in colloqui di lavoro. Tanto altri scienziati si riferiranno ad almeno uno dei parametri presentati qui per valutare il vostro 'calibro' scientifico, quando sentono parlare di voi senza conoscervi - e tanto di più se sono in qualche modo coinvolti nel vostro reclutamento, o nella vostra progressione di carriera.

1. Valutazione dello spessore scientifico di Nazioni

Riguardo a nazioni, una recente fonte di informazioni molto dettagliata viene fornita da un'istituzione spagnola – Scimago Country & Journal rank (SJR):

http://www.scimagojr.com/countryrank.php?area=0&category=0®ion=all&year=all&order=it&min=0&min_type=it

Nella classifica globale delle nazioni, l'Italia si colloca (periodo 1996-2008) all'ottavo posto, subito prima della Spagna, nel complesso delle discipline. Tuttavia, in certe discipline come Immunologia e Medicina l'Italia ha posizioni migliori, con le punte di diamante in *Haematology* (seconda) ed *Aging* (terza):

http://www.scimagojr.com/countryrank.php?area=1300&category=1302®ion=all&year=all&order=it&min=0&min_type=it

Quindi l'Italia sarebbe un buon posto per studiare l'invecchiamento. Non per niente, l'Italia ha il personale di ricerca ed universitario più anziano del mondo! In altri campi siamo anche oltre la decima posizione, soprattutto in aree di alto sviluppo tecnologico come computer sciences. Persino in certe materie umanistiche, l'Italia viene sorpassata da nazioni ben più piccole o di minore tradizione accademica. Comunque dovrebbe diventare un obiettivo fondamentale per la politica della ricerca in Italia di superare almeno il Canada nella classifica globale! Siamo il doppio dei Canadesi, e le nostre università sono molto più antiche e spesso conosciute delle loro – eppure il Canada ci batte in quasi tutte le branche del sapere.

Questo viene confermato anche dai dati sul sito In-cites: <http://in-cites.com/countries/index.html> .

2. Valutazione del prestigio scientifico di università ed istituti di ricerca

Per valutare il prestigio relativo di università ed istituti di ricerca, esistono varie fonti internazionali che compilano analisi di riferimento del loro ranking, sulla base di vari parametri spesso combinati fra di loro. Fra cui le più seguite e valide, qui segnalo:

1. ARWU (in Cina) <http://www.arwu.org/ARWU2009.jsp>
2. Times/QS (in UK) <http://www.topuniversities.com/>
3. Università di Leiden (in Olanda) http://www.cwts.nl/ranking/top250_orange_1st.html
4. Webometrics (in Spagna) <http://www.webometrics.info/top6000.asp>

Per l'Italia esiste sinora solo una 'classifica' che dia risultati consistenti ai rankings elaborati dalle agenzie internazionali 1-4: la valutazione triennale della ricerca (VTR) prodotta dal CIVR (organismo del MIUR che presto confluirà nella nuova agenzia ANVUR) per il periodo 2001-2003, VTR2001-2003:

<http://vtr2006.cineca.it/> - i cui risultati sono stati divulgati solo alla fine del 2009: http://quotidianonet.ilsole24ore.com/2009/07/24/209647-gelmini_fondi_agli_atenei_qualita.shtml

Al momento il CIVR sta lanciando un secondo esercizio di valutazione nazionale, il VQR2004-2008: http://www.civr.miur.it/modulo_eng.html.

Molto probabilmente, i risultati di questo esercizio verranno gestiti dalla nuova agenzia ANVUR.

3. Valutazione del profilo scientifico di singole persone

Il ricercatore provetto, dopo aver valutato dove (nazione ed istituto) vorrebbe andare a lavorare, bisogna che sappia valutare sia il ‘capo’ ed i colleghi che si ritroverebbe nel nuovo istituto, sia sé stesso. Aldilà della fama che certi colleghi possono avere anche solo per sentito dire, bisogna attrezzarsi con gli strumenti di valutazione quantitativa che oggi sono conosciuti da molti, e probabilmente da tutti i membri di commissioni di concorso, reclutamento e *research grants* che vi valuteranno.

Ci sono due tipi di parametri quantitativi con cui bisogna fare i conti in questo processo di auto-valutazione (come nella calibrazione di altri colleghi): 1. Valori individuali di indici come **h-index** (indice di Hirsch); 2. Valori di **Impact Factor** delle riviste in cui si pubblica. Questi parametri sono tutti derivati dall’analisi bibliometrica delle **citazioni**. Una citazione di un articolo pubblicato corrisponde alla presenza di quell’articolo nella lista delle referenze, o nel testo di *altri* articoli. Nel caso che un altro articolo sia scritto dagli stessi autori del primo articolo, si parla di *auto-citazione*. Alcuni programmi distinguono fra citazioni ‘pure’ (cioè fatte da altri scienziati) ed auto-citazioni, ma in genere vengono considerate tutte insieme, per derivare indici personali come h-index e per l’impatto di una pubblicazione. Questo perché in genere il livello di auto-citazione rimane sotto il 20% (ed il 10% viene considerato praticamente ‘fisiologico’), ed aumenta solo in certe aree di nicchia.

3.1 h-index, e derivati

L’indice di Hirsch, **h-index**, è il numero x di pubblicazioni che hanno ricevuto almeno lo stesso numero x di citazioni. Pur con i suoi limiti, il numero di Hirsch o h-index produce una valutazione quantitativa della produzione ed impatto scientifico di una persona che viene riconosciuto sempre di più. Il sito ‘Pubblico Ergo Sum’ (http://www.pubblicoergosum.org/?page_id=95) fornisce informazioni dettagliate sul significato, i limiti e le interpretazioni del h-index in Italiano. La storia del h-index in inglese si trova qui: http://www.wired.com/culture/geekipedia/magazine/17-06/mf_impactfactor?currentPage=all.

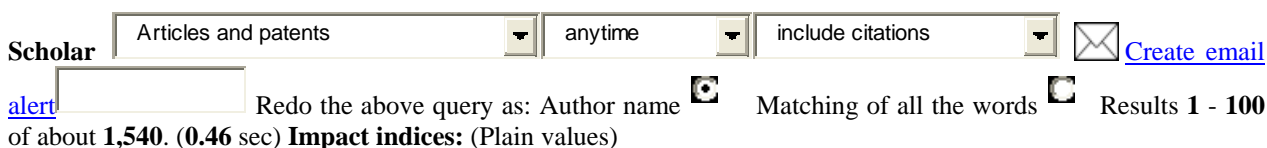
Esistono vari sistemi *gratis* per calcolare l’h-index di singole persone. Usano tutti la database di Google Scholar, oramai più espansa e dettagliata di quella ‘privata’ (cioè a pagamento) della ISI Reuters Thompson (Web of Science®, WoS) o quella di SCOPUS (Elsevier, pure a pagamento).


Il più rapido è forse questo Google gadget: <http://code.google.com/p/citations-gadget/>. Ha il vantaggio di poter calcolare anche h-index > 100, però sottostima un poco il numero di citazioni e quindi i valori di h-index, e poi non da molte opzioni per rimuovere omonimi etc. Una verifica fatta su vari accademici italiani ha dato valori simili, ma meno accurati di quelli derivati coi sistemi che ora presento qui.



3.1.1 h-index con add-on Firefox e Google Scholar

Il più potente e rapido sistema disponibile oggi è un ‘add-on’ di Mozilla Firefox che si scarica da questo sito: <https://addons.mozilla.org/en-US/firefox/addon/45283/>

Il programma, ideato dal matematico italiano GB Ianni, si apre automaticamente con la pagina di Google Scholar. Digitandovi, per esempio, il nome di P Strata, compare in cima alla schermata:



Scholar Articles and patents anytime include citations  [Create email alert](#)

Redo the above query as: Author name  Matching of all the words  Results 1 - 100 of about 1,540. (0.46 sec) **Impact indices:** (Plain values)

Citations in this page: 3181.0 **h-index** [Click to see what's this.](#): 34.0 **g-index** [Click to see what's this.](#): 50.0 **e-index** [Click to see what's this.](#): 31.0 **delta-hdelta-h: the number of citations needed to increment h-index by 1:** 4.0 **delta-gdelta-g: the number of citations needed to increment g-index by 1:** 15.0 (Normalized The citations of each paper are divided by the corresponding number of authors (authors >4 are rounded to 4) per co-authorship)

Spesso si ritrova questo messaggio in fondo al testo qui sopra:

Insufficient data in this page. Try to ask Scholar for 100 results by clicking [here](#).

Quest’opzione espande l’analisi fino al limite di h-index<100 (di default calcola solo per h-index <10) come fatto qui sopra con P Strata. Per raffinare ulteriormente la ricerca, conviene digitare il nome

completo fra doppio virgolettato, per esempio “Mauro Degli Esposti” e cliccare l’opzione ‘*Matching all the words*’. In questo modo si eliminano altri scienziati con lo stesso cognome (nel mio caso, Maria Degli-Esposti, brava collega australiana non imparentata con la mia famiglia, od un certo matematico MD Esposti, probabilmente italo-americano). Questo programma quindi offre il calcolo di 3 indici, h-, g- ed e-index e altri derivati. Di questi i più importanti sono **h-index** e **g-index**. Inoltre produce il numero totale di citazioni (in alto a sinistra: **citations in this page**) che da l’indicazione dell’impatto globale della persona ricercata – specialmente se si sono tolti gli omonimi! Il Prof. Piergiorgio Strata, per esempio, con le sue 3181 citazioni avrebbe un impatto simile al mio (con l’accorgimento qui sopra, per me si ottengono 3757 citations). Per entrambi, però il calcolo di citazioni non risulta perfettamente accurato – nel mio caso, a causa del mio cognome doppio, mancano almeno due papers con >100 citazioni ciascuno, che si estraggono solo digitando ‘M Esposti’, i cui risultati comprendono anche le citazioni degli omonimi. Questi problemi, tuttavia, producono solo una minima **sottostima** del mio h-index accurato su TUTTE le citazioni, che sarebbe 33 invece del valore di h-index = 31 ottenuto come sopra. Per scienziati della mia stessa area, biomedicina, sinora si e’ riscontrato un massimo di sottostima poco sopra il 10% (Table 1).

In pratica, in campi come la chimica, la fisica e la matematica si riduce la **sottostima** di h-index digitando solo la iniziale, invece che tutto il nome dello scienziato. Per esempio, introducendo il nome del fisico premio Nobel “Riccardo Giacconi” si ottengono 3148 citations e h-index = 16; però se si fa la ricerca digitando solo “R Giacconi”, si ottengono 7933 citations ed h-index = 39, che sicuramente e’ più vicino alla realtà. In tal modo si potrebbe incontrare il problema di omonimi scienziati che pure hanno R come iniziale. Questo si può valutare digitando solo “Giacconi” - vengono addirittura meno citazioni ed un h-index di 38; quindi in questo caso il problema di omonimi non esiste! Comunque un modo molto efficace di ridurre le omonimie e’ scegliere l’area scientifica in cui si lavora, fra le seguenti opzioni presenti nelle versioni anglosassoni (.co.uk o .com) Di Google Scholar, cliccano in alto ‘advanced Scholar search’:

- Biology, Life Sciences, and Environmental Science
- Business, Administration, Finance, and Economics
- Chemistry and Materials Science
- Engineering, Computer Science, and Mathematics
- Medicine, Pharmacology, and Veterinary Science
- Physics, Astronomy, and Planetary Science
- Social Sciences, Arts, and Humanities

Notare che il programma di add-on Mozilla ha un limite superiore, che impedisce di ottenere valori di h-index e g-index >100 – un problema decisamente minore, poiché si son ritrovati solo 2 Italiani con h-index >100 (Tabella 1 - notare come entrambi lavorino sul cancro ed in USA).

Inoltre, si noti che la lista ISI di highly cited per l’Italia non contiene tutti gli scienziati con alto h-index:

http://hcr3.isiknowledge.com/browse_author.pl?page=0&link1=Browse&valueCategory=0&valueCountry=82&submitCountry.x=16&submitCountry.y=7&submitCountry=1

Quindi, con l’aiuto di colleghi della Via-academy, ho fatto un’indagine a tappeto e compilato una lista dei ‘**top Italian scientists**’, pubblicata online dalla Via-academy in un sito dedicato, da Agosto 2010:

http://via-academy.org/Top_300_della_VIA-Academy.aspx

La tabella uno (Table 1) mostra i dati dei *top of the top* ottenuti con due metodi di calcolo del h-index. Quello ‘add-on Firefox’ menzionato qui sopra, e quello del ISI Web of Science® (WoS) presentato subito dopo (parte 3.1.2). In tutti i casi, incluso quelli rari di **inusuale sovrastima** (come per alcuni fisici in Table 3 qui sotto), ho consistentemente seguito questo approccio per discriminare l’inserimento di una persona nella tabella dei ‘top Italian scientists’ http://www.topitalianscientists.org/top_italian_scientists.aspx :

per entrare nella lista bisogna che i valori di h-index siano >30 in almeno 2 dei metodi descritti in questo manualetto (3.1.1, 3.1.2 o 3.1.3). In genere si inserisce il valore maggiore, a meno di una forte presenza di *citazioni minori* o di sottostime sistematiche (vedi sotto).

Table 1 – h-index for top Italian biomedical scientists with ISI Web of Science® and Google Scholar

Scientist	Google Scholar h-index+	ISI WoS h-index*	location
Carlo Croce	121	124	USA
Napoleone Ferrara	103	106	USA
Giorgio Trinchieri	98	114	USA/France
Alberto Mantovani	97	102	Milano
Antonio Lanzavecchia	94	92	Switzerland
Alex Sette	94	96	USA
Alessandro Moretta	92	88	Genova
Giuseppe Remuzzi	90	91	Milano
Giuseppe Mancina	86	90	Milano Bicocca
Lorenzo Moretta	86	95	Genova
Pierpaolo Pandolfi	86	81	USA
Pietro De Camilli	84	56	USA
Riccardo Dalla Favera	84	48	USA
Pier Giuseppe Pelicci	83	89	Milano
Silvia Franceschi	83	81	France/Torino

+including all entries under the “full name” using Firefox add-on; *limited to articles and reviews.

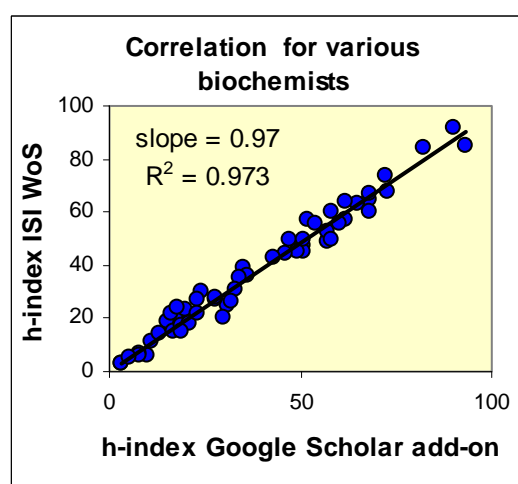
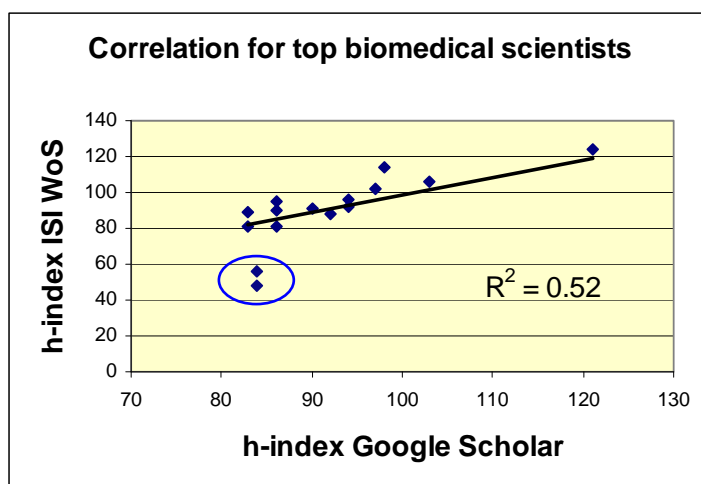


Figure 1 – Correlation between the h-index obtained with ISI WoS and the h-index calculated with the add-on Firefox for Google Scholar (standard method for producing the list of ‘top Italian scientists’). **A.** Results were obtained with the same set of scientists as in Table 1. Although the overall correlation is poor, the results fit a 1:1 correspondence (slope = 0.98). The two outliers in the blue circle, under-estimated by ISI WoS, are *much younger* than all the others and have a *composite name*: De Camilli and Dalla Favera. Taking these scientists out, the correlation becomes much better, with $R^2 = 0.80$ (slope = 1.03). Adding then the h-index values of several other biomedical scientists, including those examined later in Fig. 2 and myself, the slope approaches 1.0, as shown in part B. In general, age and name complexity can alter the correspondence between Google Scholar and ISI values of h-index, but in the opposite direction. The *older* the scientists, the *lower* will be the values for Google Scholar vs. ISI; indeed, Google Scholar has a poor coverage of papers published earlier than the nineties. *The opposite can also be true.* Consequently, Google Scholar will provide a more comprehensive coverage than ISI WoS for young or middle-age scientists. Moreover, ISI WoS does not recognize well composite surnames starting with ‘Di, De, Da, Dalla, Della, Degli’, which are relatively common among Italian scientists. Google Scholar has a more powerful search algorithm for finding these surnames more precisely than ISI. **B.** Results were obtained with a set of 54 biochemists working in Italy and have been already published online here: <http://www.lavoce.info/articoli/pagina1001893.html> .

3.1.2 h-index con ISI Web of Sciences®

Allora, come si calcola il valore di h-index con WoS? Il sistema risulta molto più complicato di quello presentato qui sopra, ma sinora è stato considerato quello più ‘ufficiale’ da molti scienziati ed istituzioni, dato il dominio che il database ISI ha avuto nel campo dell’analisi bibliometrica. Le istruzioni sono le seguenti. Innanzitutto si va al sito ufficiale del WoS (dentro a quello più generale di Web of Knowledge Thomson Reuter), cui si accede tramite la sottoscrizione dell’ente di appartenenza (quasi tutte le università e centri di ricerca pagano per averlo e metterlo a disposizione del loro personale):

http://apps.isiknowledge.com/WOS_GeneralSearch_input.do?product=WOS&search_mode=GeneralSearch&SID=R1277f5JKj6hL8F9Pd4&preferencesSaved=

Poi si va alla opzione ‘Search for Author’ (seconda linea), per esempio (nome preso a caso):

Mancini S* per una scienziata chiamata Stefania Mancini (un nome reale, come si vedrà).

Prima di effettuare la ricerca col tasto SEARCH conviene settare le opzioni come indicato qui sotto:

Timespan:

All Years (updated 2010-07-31)

From 1945-1954 to 2010 (default is all years)

Citation Databases:

Science Citation Index Expanded (SCI-EXPANDED) --1945-present

Social Sciences Citation Index (SSCI) --1956-present

Arts & Humanities Citation Index (A&HCI) --1975-present

NEW! Conference Proceedings Citation Index- Science (CPCI-S) --1990-present

NEW! Conference Proceedings Citation Index- Social Science & Humanities (CPCI-SSH) --1990-present

Per scienziati non di *Social and Humanity areas*, si lascia solo il tick al data base superiore. Poniamo che la Mancini sia attiva dal 1980, allora l’anno lo si digita nel Timespan:

Timespan:

All Years (updated 2010-07-31)

From 1980 to 2010 (default is all years)

Con queste opzioni si dicavano 324 *results* (cioè papers o proceedings). Per rimuovere omonimi (e ci sono vari S Mancini oltre a Stefania!), conviene selezionare solo i risultati del campo specifico in cui Stefania lavora di più (diciamo, *immunology/hematology*), che si fa mettendo un *tick* a questi campi selezionati in basso a sinistra dal programma, e poi si clicca: REFINE. Ora rimangono solo 30 pubblicazioni. Per ricavarsi il valore di h-index da queste pubblicazioni si clicca il tasto nella parte destra in alto, con un simbolo a grafichetto denominato: ‘create citation report’. Dopo poco appaiono due grafici sulla schermata ed un a colonna a destra, sul cui fondo compare il valore di **h-index** per le pubblicazioni selezionate sopra:

Results found: 30

Sum of the Times Cited [?]: 182

[View Citing Articles](#)
[View without self-citations](#)

Average Citations per Item [?]
 : 6.07

h-index [?]: 6

So, Stefania ha published six papers that have received at least 6 citations so far!

Questo è un valore normale per una scienziata all'inizio della sua carriera, ma un po' limitato considerando che Stefania probabilmente lavora da quasi 30 anni. In realtà Stefania ha pubblicato papers anche in altre aree scientifiche come Oncologia, e quindi la ricerca va rifatta in modo meno selettivo, comprendendo questa ed altre aree bio-mediche, ma tralasciando le pubblicazioni in aree non congruenti col suo profilo di scienziata, per esempio fisica delle particelle, usando le opzioni indicate su questa pagina: http://images.isiknowledge.com/WOKRS49B3/help/WOS/h_da_sets.html.

Rifacendo la procedura suddetta, si ottengono così 51 papers e **h-index = 11**.

3.1.3 h-index con PorP

Molto conosciuto risulta poi il programma di Anne Harzing, che si scarica gratis dal sito 'Publish or Perish' <http://www.harzing.com/pop.htm>

Il programma, generalmente abbreviato 'PorP', usa anch'esso Google Scholar come database, ma ha una serie di opzioni di selezione ed outputs maggiormente variegata. In particolare si possono selezionare i papers dell'autore (opzione utile per rimuovere omonimi) e si ottengono anche questi parametri:

Total number of papers, Total number of citations, Average number of citations per paper, Average number of citations per author. Average number of papers per author etc.

Alcuni parametri sono utili, ma altri sono dei semplici esercizi accademici. Alla fine dei conti, tutti gli indici sono derivati dal **classico h-index** di Hirsh, che rimane il più usato e conosciuto! Seguito forse dal g-index di Egge, che è sempre maggiore dell' h-index e valuta meglio i papers che ricevono molte citazioni. Il sito della Harzing comprende anche delle valutazioni tecniche ed accademiche dei pregi e limiti di questi indici in questa rassegna online: *Reflections on the h-index*, Prof. Anne-Wil Harzing, University of Melbourne, ©Copyright 2007-2008 Anne-Wil Harzing. Third version, 23 April 2008, http://www.harzing.com/pop_hindex.htm.

Nel caso della nostra Stefania Mancini usata come esempio prima per WoS, con PorP si ottengono, dopo aver selezionato le aree *Biology & Life Sciences e Medicine* (cliccando in alto a destra) e digitando 1980 – 2010 come periodo: 51 papers e **h-index = 14**. I dati sono congruenti, ma il valore di h-index risulta maggiore che con l'analisi WoS, come spesso succede anche con add-on Firefox. Tuttavia la correlazione fra PorP e ISI sembra piuttosto buona per scienziati biomedici ed altre disciplines 'hard' (Fig. 2).

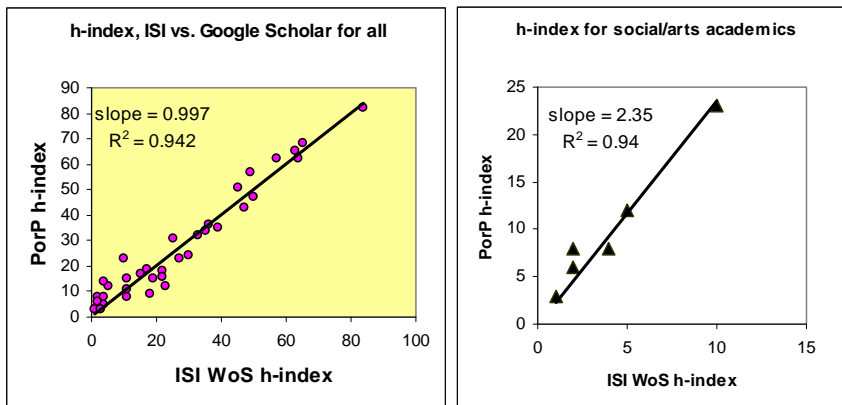


Figure 2 – Correlation of the h-index calculated using PorP and that evaluated with ISI WoS. **Left**, data for a set of 40 Italian scientists, predominantly of biomedicine including several from Fig. 1 and all the potential candidates for ANVUR: <http://rpc264.cs.man.ac.uk/VIA/index.php/ANVUR>. **Right**, different correlation for 7 for scholars of social sciences and arts from the set in the left. In this subset, the data fits a slope clearly larger than as for all scientists together (left). Note that a very limited number of these scholars could enter the list of 'top Italian scientists':

http://www.topitalianscientists.org/Top_300_italian_scientists_VIA-Academy.aspx.

Per esperienza, il programma PorP risulta decisamente lento rispetto a programmi inseriti come l'add-on Firefox descritto sopra (3.1.1) ed ha un limite superiore di 1000 *entries*, che non permette una analisi completa e dettagliata di scienziati molto produttivi. Comunque, globalmente i valori di h-index ricavati con questo programma correlano molto bene con quelli ricavati con ISI WoS, eccetto per studiosi ed accademici di economia, scienze sociali, *arts and humanities* (Fig. 2). Google Scholar ha una database più ampia e dettagliata di quella prodotta da ISI WoS, come ben documentato nel sito della Harzing (<http://www.harzing.com/pop.htm>), dove si sottolinea la validità, oltre che la facile accessibilità, delle ricerche bibliometriche fatte *bene* con Google Scholar!

3.1.4 Correlazioni e discrepanze nei dati bibliometrici ottenuti con Google Scholar e ISI WoS

Diversi metodi spesso concordano fra loro nel definire numericamente la stima del valore di h-index, ma talvolta discordano, anche in maniera significativa. La correlazione fra i valori di h-index ottenuti col metodo 3.1.1 (add-on Firefox) ed il metodo 3.1.3 (PorP Harzing) è consistentemente alta, come riportato qui sopra, anche se il primo produce meno sottostime del secondo (3% in media per biomedici). Le maggiori differenze si riscontrano fra scienziati multidisciplinari e certi economisti che fanno molto 'networking', cioè attirano un volume di *citazioni minori* maggiore del normale. Il database di Google Scholar, che è il più ampio fra quelli disponibili, contiene anche citazioni di libri e di capitoli di libri, insieme a queste *citazioni minori*: a) 'editorship' (a cura di) di libri e proceedings di congressi, simposi etc.; b) 'acknowledgements' (ringraziamenti); c) websites ed online repositories; d) tesi; e) reprints online; f) abstracts (di congressi etc.). PorP filtra la maggior parte di queste citazioni mentre add-on Firefox no. Il database ISI Web of Science® (WoS) poi deriva dal classico Citation index, il quale sin dall'inizio era ben collegato con il più vecchio database esistente in scienza, quello dei **Chemical Abstracts** http://en.wikipedia.org/wiki/Chemical_Abstracts_Service, e quindi copre bene pubblicazioni di discipline chimiche. Tuttavia non cataloga libri e neppure a)-f); le citazioni sono conseguentemente limitate a quelle che appaiono nella lista delle referenze negli articoli ed altri 'editorial materials' (incluso 'corrections') contenuti nelle riviste scientifiche, anche quelle molto vecchie che ora son scomparse, o non legge più nessuno. Inoltre ISI non lista riviste minori o molto recenti, mentre esclude quelle in lingue diverse dall'inglese. Quindi la dimensione delle citazioni, su cui si basa il calcolo di h-index (e di altri parametri bibliometrici come l'Impact Factor, vedi parte 3.2), varia se si usa Google Scholar o ISI WoS, specialmente per alcune discipline come la chimica, mentre è comparabile per le discipline mediche, biologiche e fisiche - che in effetti comprendono la maggioranza dei top Italian scientists.

Nei riguardi degli studiosi italiani di discipline non *hard sciences* e non biomediche, la copertura di ISI WoS risulta molto limitata rispetto a Google Scholar. Per dare esempi eclatanti, ho comparato l'h-index ottenuto con Google Scholar con quello di ISI WoS per alcune persone conosciute (Table 2).

Table 2 – Calculated h-index for Italian scholars with ISI Web of Science® and Google Scholar

Scholars	Google Scholar h-index+	ISI WoS h-index*	Notes
Umberto Eco	68	6	Semeiotica, scrittore
F Modigliani	63	25	Nobel, Economista, USA
G Tabellini	62	22	Economista
Tito Boeri	42	11	Economista/giornalista
F Giavazzi	43	9	Economista/giornalista
G Calabresi	42	13	Studi giuridici, USA
Giovanni Sartori	36	14	Scienze politiche
Dario Fo'	20	2	Nobel, teatro
Dacia Maraini	8	1	Scrittrice, teatro
Italo Calvino	33	1	Scrittore

+including all entries under the "full name" using Firefox add-on; *over all 3 databases, no limits.

Chiaramente, Google Scholar produce valori di h-index decisamente maggiori che quelli ottenuti con il database ISI WoS, confermando in modo inequivocabile le conclusioni precedentemente raggiunte dalla Harzing ed altri studiosi stranieri: <http://www.harzing.com/pop.htm> .

Nel mondo ideale bisognerebbe produrre un database integrato con Google Scholar che contenesse quanti più lavori e pubblicazioni di ogni genere in italiano – un bel progetto per informatici ed umanisti!

Ma torniamo alle scienze, incluso quelle economiche, che per approccio e metodologia si avvicinano maggiormente, fra le discipline sociali, a quelle *hard*. La Tabella 3 (**Table 3**) mette insieme rappresentanti di queste discipline con i valori di h-index ricavati con ciascuno dei 3 metodi descritti. Comprende anche un gruppo di economisti italiani selezionato fra quelli considerati ‘top’ nelle discipline economiche.

Table 3 – Values of h-index calculated with 3 different methods for scientists in the Via-academy list

Area and scientist	Google Scholar h-index+	ISI WoS h-index*	PorP h-index£	notes
<i>Physics</i>				
G Martinelli	57	56	56	
A Sagnotti	39	31	39	
Roberto Pittau	31	20	26	
M Giaquinta	33	34	20	
S Frixione	41	25	35	
P Gambino	39	28	36	
<i>Chemistry</i>				
V Balzani	59	76	58	ISI value
M Parrinello	73	85	73	
V Barone	53	59	53	ISI value
D Braga	33	47	25	ISI value
M Prato	53	61	53	ISI value
F Cacialli	34	38	23	ISI value
C Taliani	32	35	16	ISI value
G Pacchioni	48	57	49	ISI value
<i>Economics[%]</i>				
Alberto Alesina	88	34	87	
Marco Pagano	50	21	47	
Roberto Perotti	43	16	41	
Orazio Attanasio	41	16	40	
Tullio Jappelli	37	15	35	
Luigi Guiso	36	14	33	

+including the entries under the initial only in the area of Physics and Astrophysics (except Giaquinta, who also publishes in the area of mathematics) using Firefox add-on – values **red** are the numbers inserted after this analysis in the Via-academy list, especially in the updated version with disciplines posted on 21 Sept 2010 (notes);

*With initial, restricted to papers and review only in physics sectors, all years;

£With initial, restricted to the area of Physics, all years;

#Number of papers and books listed in the first page of the Firefox add-on with max selection of 100;

%Queries with “full names” using add-on Firefox (all fields) and PorP (Business and Social) but initial for ISI WoS.

I dati nella Tabella 3 suggeriscono che i chimici sono sistematicamente penalizzati da Google Scholar rispetto a ISI WoS. Tuttavia l'opposto succede per gli economisti (Tabella 3 e Fig. 3). La fondamentale ragione della sottostima del valore di h-index per i chimici risiede nella incompletezza del database Google Scholar per pubblicazioni di chimica, soprattutto quelle vecchie che sono invece minuziosamente catalogate in ISI WoS, sin dall'origine ben integrato coi CHEMICAL ABSTRACTS. Quindi il database ISI e' accuratissimo per i chimici, anche perchè contiene riviste oramai desuete, ma strapazza gli economisti (Fig. 3), come pure gli astrofisici, i computer scientists ed i matematici. Per non parlare degli umanisti (Table 2). Inoltre perde colpi rispetto a Google Scholar per tutte le branche della biologia moderna. Quindi Google Scholar è il grande 'omogenizzatore' democratico per tutte le scienze - eccetto che per chimica, dove ISI rimane il 'vecchio giustiziere'. Per i fisici, sembra che ISI WoS e Google Scholar siano più o meno simili (Table 3); infatti ci sono forse state più lamentele di sovrastima che di sottostima per i valori degli scienziati di fisica riportati nella lista della Via-academy. Notare infine come in quasi tutti i casi (Tabelle 1, 2 e 3) il metodo add-on Firefox dia valori intermedi fra quelli maggiormente sottostimati e quelli più completi di ISI per i chimici – e spesso leggermente maggiori di quelli di PorP.

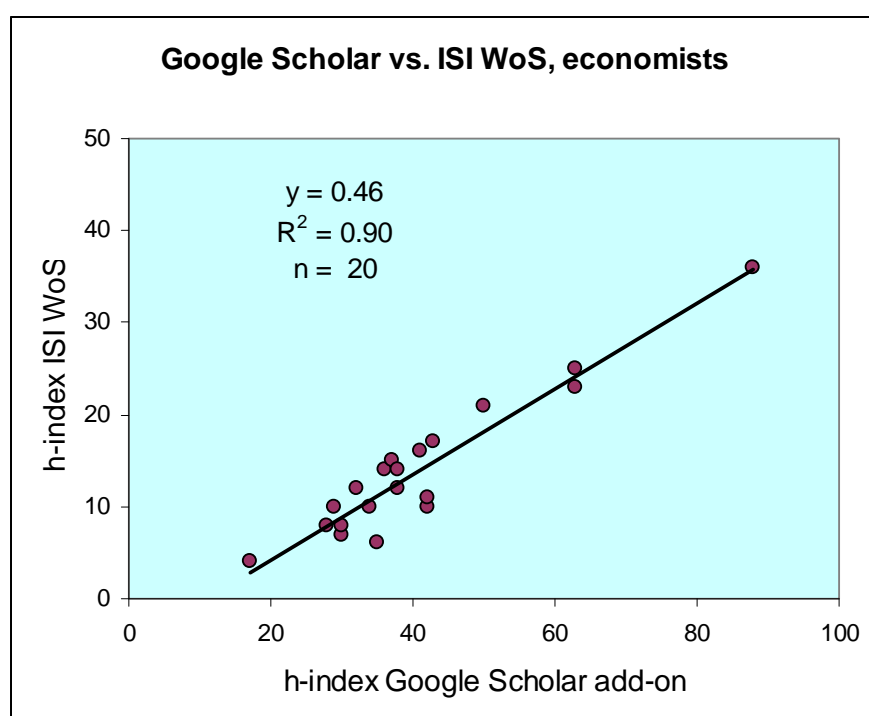


Figure 3 – Correlation of the h-index values obtained with Google Scholar and ISI for Italian economists. Twenty economists, including most of those present in the list of 'top Italian scientists', were evaluated using the specific areas of 'Business' and 'Social Sciences' in the searches with either Google Scholar (add-on method 3.1.1) or ISI WoS (3.1.2). Note that the slope is approx. 0.5, i.e. ISI under-estimates the h-index values of economists by, on average, a factor of two.

3.2 Impact Factor – IF

L'altro parametro fondamentale per la valutazione quantitative dei singoli è il fattore di impatto, **IF**. Metto qui la sua definizione in inglese, come ufficialmente riportato da chi lo ha inventato e lo continua a commercializzare - in inglese, *of course!*

The concept of IF has been developed by the Institute of Scientific Investigation (ISI) in USA, now part of the Thompson Reuters group, since the '60', using the data in the publications of Journal Citation Reports® (JCR®): http://thomsonreuters.com/products_services/science/free/essays/impact_factor/ The impact factor is a measure of the frequency with which the "average article" in a journal has been cited in a particular year. The annual JCR impact factor is a ratio between citations and recent citable items published. Thus, the impact factor of a journal is calculated by dividing the number of current year citations to the source items published in that journal during the previous two years. It changes slightly every year, but for most journals is now fixed in a relatively narrow range of values.

Per una rivista tipo di biologia cellulare, *Journal of Cell Biology*, il valore di IF rimane intorno a 10 da molti anni. Questo significa che, in media, un articolo pubblicato su *Journal of Cell Biology* riceve 10 citazioni all'anno. Chiaramente, non tutti gli anni lo stesso numero, e non tutti gli articoli si beccano 10 citazioni annuali. Per esempio, il mio unico articolo su questo giornale, pubblicato nel 1999, ne ha ricevuto 221 sinora - quindi: $220/11 = 20$ citazioni all'anno. Chiaramente questo articolo alza la media del giornale, dove ce ne saranno altri che si beccano anche solo 5 citazioni all'anno. Ma la media sarà, appunto, intorno a 10! Inoltre è praticamente impossibile che questa rivista pubblichi papers che ricevano pochissime citazioni. Il prestigio della rivista e del suo comitato editoriale è troppo alto per far sì che questo possa accadere con frequenze superiori al rumore di fondo. Ogni campo del sapere ha le sue riviste top, che attirano un numero medio di citazioni che dipende dall'ampiezza del campo e dai soldi che vengono investiti nello stesso campo. Per questo riviste mediche, anche ad indirizzo clinico, si prendono molte più citazioni che riviste di chimica; per esempio, *Diabetes* ha un IF almeno il doppio di quello di *JACS*, la rivista *top of the top* in chimica! Comunque, nel mondo scientifico attuale dominato dalle aree biomediche, le valutazioni bibliometriche di IF vanno calibrate per la diversa frequenza di citazioni che caratterizza i vari campi del sapere, e conseguentemente si riflette nel IF delle riviste specializzate nei vari settori. Recentemente, analisi basate sull'algoritmo Page (usate da Google, tanto per cambiare) aiutano a ricalibrare l'IF relativo delle riviste nei vari campi del sapere, come viene ora sistematicamente analizzato nel sito SJR (vedasi la parte iniziale, **1** di questo *manualetto*): http://www.scimagojr.com/countryrank.php?area=0&category=0®ion=all&year=all&order=it&min=0&min_type=it

In ogni caso, i ricercatori provetti si prenderanno cura di calcolare l'**impatto effettivo** delle loro pubblicazioni, che si ricava dalle citazioni riportate per ogni loro paper da Google Scholar o ISI WoS con i metodi descritti qui sopra. Se poi nel concorso si richiede esplicitamente di riportare l'IF delle riviste su cui si abbia pubblicato, i ricercatori provetti valuteranno il loro IF 'stimato' dai valori ufficiali che ISI assegna ogni anno a ciascuna rivista, andando al sito WoS/Knowledge e ricavandosi i valori corrispondenti. Molto spesso, tuttavia, sarà sufficiente scandagliare il web per i valori di IF delle riviste del campo, che magari saranno dell'anno precedente (certamente disponibili sono gli IF per quasi tutte le riviste dell'anno 2008), ma tanto cambiano ben poco di anno in anno. Si prenderanno inoltre cura di annotare sul loro CV che i calcoli eseguiti sono stati fatti *tutti* sulla base dei valori IF del 2008 come riportati sul sito usato, e debitamente citato. Questo vale anche per i valori di h-index od altri parametri.

*Ultimately, the key parameter remains the number of **citations** that your work has attracted in your field. The aim of your career is to keep continuity in your productivity and thus consolidate your scientific impact!*

Mauro Degli Esposti