How Scholarly is your University? Ask Google Scholar

Ray K Iles

A lot of money is paid for publicity agencies and in-house staff to promote the merits of any given university or Higher Education institution, but how much is hype, historical, or downright misleading? One beacon from which the dark shadows recede is the UK National Student Survey, a hard metric which will soon be burning a branding mark on all UK University degrees. Not all will display the mark with pride, as they advertise their wares in the national and international bazaars of student recruitment.

But what about scholarship? After all, it is not all about students' perceptions of whether they had a good time. Indeed, the academic standing, scholarly works and impact of the research of a University is a vexing question for governments, who fund and regulate HEIs. It is also for academic staff and the serious potential undergraduate or graduate student – so how much of a master of the arts & sciences is your potential *Alma Mater?*

Such is the magnitude of the question that, every 5-7 years, in the UK we have an audit of all HEIs for research excellence (RAE/REF) and, this time, impact is a major driver in the evaluation exercise. For the UK government this will be the rock upon which to base the distribution of funds for research. For overseas governments, it is often used as the criterion upon which to decide sponsorship or non-sponsorship of a student's studies in the UK. For newspapers, this will be one hard metric which remains constant in the annual league tables. The universities themselves will use this score - or more likely, the various descriptive that results - for years, moulding the resulting metrics to whatever illusion of publicity they can.

But how hard a metric is the RAE/REF? It is always going to be historic and, as time passes, dated. After 5 to 7 years, it will be so dated that it bears no resemblance to the current situation. Furthermore, there will always be the criticism that the score awarded is a subjective evaluation plagued by bias (however subconscious) in favour of historic institutions.

The research evaluation framework for 2013 has tried to move away from the criticism and suspicion that accompanies expert opinion by introducing citation metrics. Unfortunately, the adoption of citation data in the process has not been universally accepted, and will only be available to a restricted number of units of assessment, notably traditional science, and not arts, humanities and social sciences.

However, this Canute-like stand in 2013 looks very likely to be washed away by a tidal-wave of bibliometric data, and washed away very quickly indeed by Open Access, and the rise of academic social networks. Sites such as LinkedIn and ResearchGate are beginning to be used by academics to voice questions of concern about the direction of academia, as much as they are about developing scientific knowledge and exchange of data. It is the perceived acceptance of bibliometric data voiced in ResearchGate that makes me believe such a seismic change has occurred: I was introduced to

bibliometrics twenty five years ago with my medical school's adoption of the Thompson ISI Web of Knowledge. Even then, you could look at how many times a paper had been cited, but it was limited to ISI listed publications. It was often rejected as a valid measure of what we now call 'impact' because of self-citation, but that argument has now largely been ignored – if you get published, so what if you cite your previous work? Several other major attitudinal changes have occurred since then. As detailed in a discussion thread I have contributed to in ResearchGate, what is more important in evaluating an individual's academic standing and performance - the impact factor of the journal they publish in, or the number of citations the work itself receives? The overwhelming conclusion has been how many times the work is cited. This chimes with other developments in concepts of evaluation measures of academia. Productivity might be how much and how often you publish, but the impact value of one's work is borne out in its citation.

The Hirsch index, marries productivity with impact. It looks at the outputs (papers/books or chapters, etc.) published by an academic to date, and ranks them in order of the highest citation to the lowest. The h-index score is the rank order number of published outputs that equals or exceeds the number of citations. This therefore balances volume of output with citation. There are various proposed modifications of this, but the h-index is accepted by the majority.

The other major force has been the development of online journals, and search engine capacity for recording and collating citation by Thompson ISI, Elsevier Scopus and Google Scholar. Google Scholar provides a citation indexing that surveys far more published material than either of the other systems. Of equal importance, and why Google Scholar has such a global following, is that unlike the other two, Google Scholar is free.

In the last couple of years Google Scholar introduced a feature which invited academics who use it to register and identify their captured publications. After validating their authorships/co-authorship, Google Scholar calculates their total citation, h-index and also an i10 index - how many papers have been cited 10 or more times. Furthermore, for current rather than total historic evaluation, Google Scholar also calculates the total citation, h-index and i10 for citations of an author's published work in the last 5 years.

Total citations number, H-index and i10 index are rapidly becoming the metrics with which to evaluate your career, and your standing as a scholar. Of course it requires integrity of the academic; after all, there are multiple authors with the surname Ahmed, Jones, Patel, Smith, Shah, Wu etc., and initials are not always fully listed, nor useful. Thomson ISI and Scopus are promoting unique researcher ID to aid in their validation of academic ownership, but as more and more academics identify their published outputs, any mis-assignment of work will be discovered and resolved.

To come back to my original question, can Google scholar be used to evaluate how scholarly a university is? Well, yes, it can, because of what really constitutes a 'university'. It is not a group of buildings, it is a collective of academics who teach and research in those building, and it is a dynamic entity. If you type in a university name in the Google Scholar Citation author box, it will list all academics who have registered and identify themselves as part of that institution. There are a few confusions, but you can eliminate all whose listed email address does not correspond with that institution, thus giving you confidence in the data. Furthermore, not only did it eliminate those who have adjunct positions, but also allows one to exclude 'the great and good' who were unfortunately now deceased. Isaac Newton and John Maynard Keynes had seemingly identified themselves as being professors at Cambridge University, and apparently validated their ownership of work from beyond the grave (email addresses withheld).

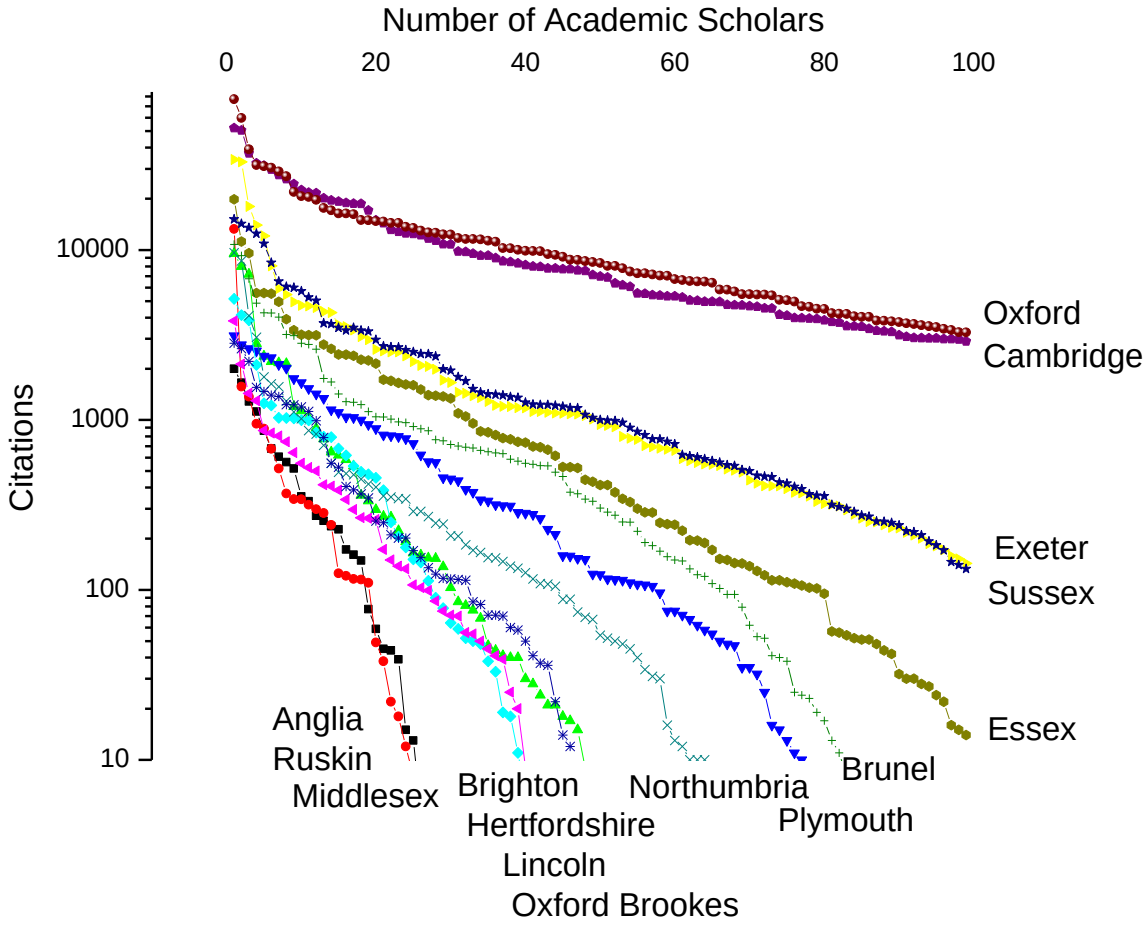

Figure 1. – Hirsch type plot of Google Scholars registered at selected example UK Universities ranked in order of total number of times their published work had been cited, highest to lowest (minimum 10) maximum of top 100 scholars shown. Data collected May 2013.

Nevertheless, this is a very rich set of bibliometrics from which to make a judgement of an institution's scholarship. I have plotted the total citation of institutions' self-identified academics in rank order, in the same way you do in the Hirsch index, and found the HEI Scholarship H-index of a number of universities as shown in the figure and table. As detailed in the table, from such data I have also adopted a scaled up a version of the i10 index: the HEI s-1000 score, i.e. how many of the institutions' academics who are registered have been cited more than a 1000 times. The ranking is revealing, and it will be interesting to compare the results of the REF2013 with a ranking based on my HEI Scholarship s-1000 index and H-index.

Rank	Institution	HEI s-1000 index	HEI scholarship H-index
1	University College London	272	399
2	Manchester University	201	328
3	Oxford	210	317
4	Imperial College London	186	309
5	Cambridge	183	318
6	Edinburgh	156	271
7	Southampton	143	233
8	Sheffield	118	220
9	Kings College London	118	212
10	York	114	205
11	Glasgow	109	203
12	Loughborough	108	234
13	Bristol	108	213
14	Birmingham	101	195
15	London School of Economic	101	192
16	Liverpool	100	190
17	Leeds	99	187
18	Nottingham	98	172
19	Aberdeen	97	181
20	Queen Mary	96	173
21	Warwick	94	179
22	Durham	79	146
23	Bath	76	143
24	Newcastle	66	135
25	Exeter	60	123
26	St Andrews	56	101
27	Sussex	55	113
28	Cardiff	54	116
29	Lancaster	53	101
30	Sussex	51	103
31	Surrey	50	116
32	Royal Holloway	48	75
33	Leicester	47	93
34	Bangor	44	97
35	Reading	43	112
36	East Anglia	43	107
37	Kent	43	106
38	Strathclyde	41	101
39	Queens, Belfast	39	114
40	Essex	38	91
41	Swansea	36	92
42	Dundee	33	70
43	Brunel	30	89
44	Stirling	30	77
45	City	29	85
46	Aberystwyth	27	63
47	Plymouth	26	74
48	Heriot-Watt	24	73

49	Aston	23	64
50	Lincoln	18	60
51	Hull	18	57
52	Hertfordshire	18	41
53	Portsmouth	17	55
-	West of England	17	55
55	Goldsmiths UoL	14	31
56	Oxford Brookes	12	45
57	Keele	12	34
58	Northumbria	11	58
59	Manchester Metropolitan	11	45
60	Ulster	11	43
61	Central Lancashire	11	35
62	Greenwich	10	41
63	Kingston	9	32
64	Salford	8	40
65	Edinburgh Napier	8	22
66	Glasgow Caledonian	7	42
67	Bradford	7	29
68	Nottingham Trent	6	48
69	Roehampton	6	33
70	Liverpool John Moores	6	25
71	De Monfort	5	37
72	Bournemouth	5	29
73	Brighton	5	38
74	St Georges, University of London	5	13
75	Middlesex	4	33
76	Teesside	4	27
77	Anglia Ruskin	4	23
78	Robert Gordon	4	20
-	Bedfordshire	4	20
80	Glamorgan	4	15
81	Coventry	3	33
82	Huddersfield	3	24
83	Derby	3	21
84	Sheffield Hallam	2	26
85	Gloucestershire	2	23
86	Leeds Metropolitan	2	13
87	Northampton	2	11
-	Sunderland	2	11
89	Westminster	2	10
90	Cardiff Metropolitan	2	9
91	Buckingham	2	6
92	Birmingham City	1	23
93	Chester	1	9
94	Worcester	1	7
95	Cumbria	1	6
96	Winchester	1	5
-	Queen Margaret	1	5
98	Abertay Dundee	1	4

99	Bath Spa	0	9
100	Edge Hill	0	8
-	Staffordshire	0	8
-	Canterbury Christ Church	0	8
103	Harper Adams	0	4
	University of the Arts, London	0	4
105	Southampton Solent	0	3
-	Falmouth	0	3
-	Buckinghamshire New	0	3
-	York St John	0	3
109	Chichester	0	2
110	University for the creative arts	0	1
111	Arts University at Bournemouth	0	0
-	Newman	0	0
-	Leeds Trinity	0	0

Bibliometrics scores derived from Google scholar citation for UK higher education institutions on the 10th June 2013.

Of course, all of this does currently depend on how switched on to bibliometrics and the pulse of opinion to the world community of research an individual academic and institution are. But is that not the point? UCL, Manchester, Imperial, Cambridge, Oxford, and others in the UK Ivy league, the Russell group, are clearly in the highest echelons; and a comparison of Google Scholar bibliometrics for the world top ranking universities is warranted. Conversely the contrast of the top 20% of UK HEI's with the bottom 20% (even 40%) is so stark that one might questions what the definition of a UK University is (see figure 2) or do they simply not know how to play the game?

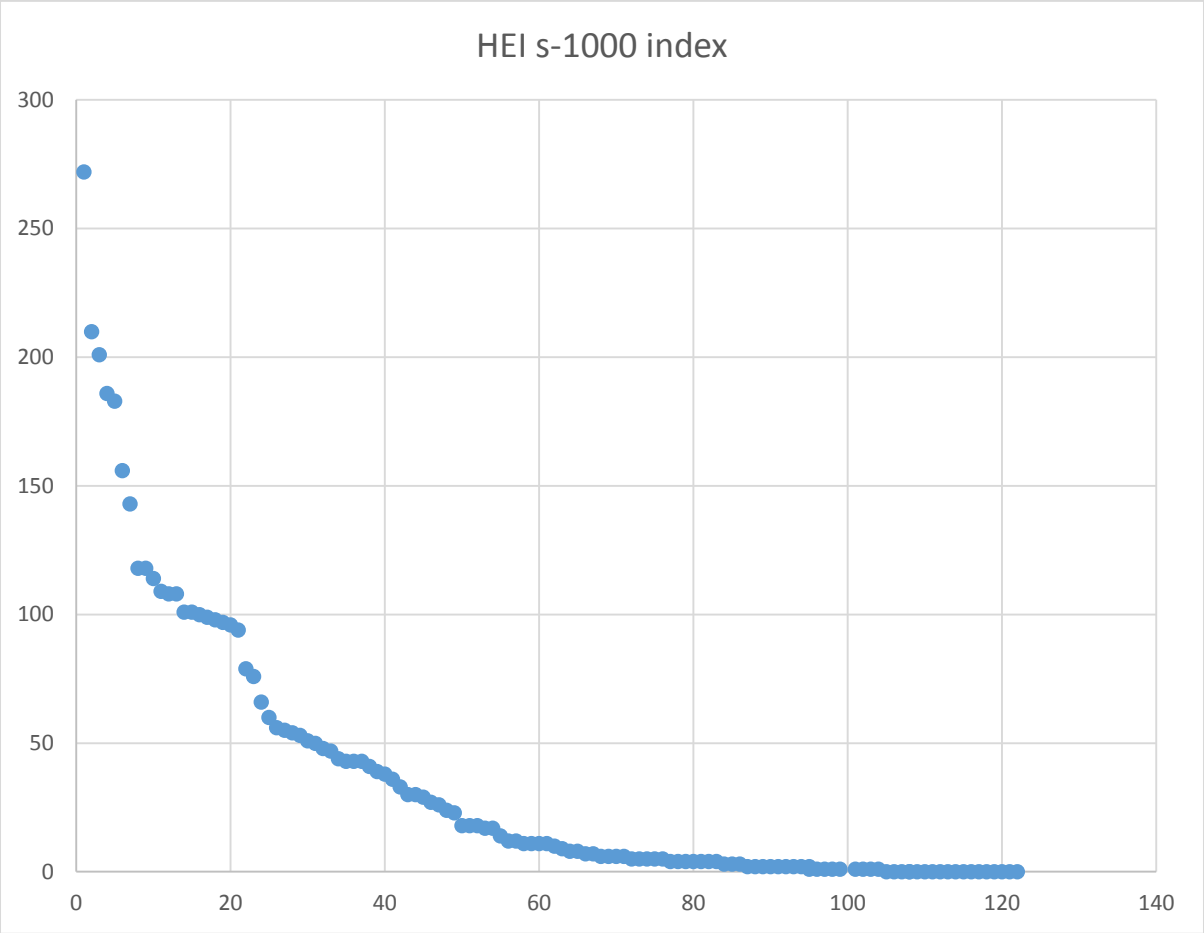

Figure 2. Plot of number of academics registered on Google Scholars whose work has been cited more than a 1000 times and are associated with a University/HEI (s-1000 score of a university) plotted against the rank order of that University/HEI according to the s-1000 score.

Google Scholar Citation should not be such a surprise if you are internationally competitive, and it speaks volumes if it does.