3-1-2009

Time is Money, More Bang for the Buck: An Introduction of the GC-Index for Assessing Research

Hirotoshi Takeda
htakeda@cis.gsu.edu

TIME IS MONEY, MORE BANG FOR THE BUCK: AN INTRODUCTION OF THE GC-INDEX FOR ASSESSING RESEARCH

Hirotoshi Takeda
Georgia State University
htakeda@cis.gsu.edu
University of Paris Dauphine
hirotoshi.takeda.05@campus.dauphine.fr

ABSTRACT

The current research expands on past research in ranking of information systems journals. A new metric, the gc-index, derived from previous Hirsch family of bibliometric indices, is introduced. Past indices didn't take into account the time or size of citations or ignored both. The gc-index expands on the h-family of indices by taking into account both the time the research has been in existence, via year, and the amount of impact, via citation count. The metric is used to update IS journal rankings. The metric is compared with the previous indices that take into account the size (g-index) and time (hc-index).

Keywords

Citation, citation analysis, citation count, bibliometric tools, h-index, g-index, hc-index.

INTRODUCTION

Previous work in this stream of papers have appeared in the proceedings of the 2008 SAIS conference, 2008 AMCIS conference, and 2008 ICIS conference. This paper adds the new metric of the gc-index, which has not been introduced in any outlet, as far as the literature review by the author is concerned. The new index can be used in other discipline areas outside of IS and is considered to be significant enough to warrant a submission to SAIS 2009.

JOURNAL/RESEARCHER RANKING HISTORY

Journal and researcher ranking studies have been conducted in order to assess the worthiness of a journal or a researcher (Alexander, Sherer, and Lecoutre, 2007; Lowery, Karuga, and Richardson, 2007; Rainer, Kelly, and Miller, 2005). These ranking studies are trying to justify where the worth of the field lies. Whether the worth be in a particular journal or a particular researcher. In turn, these ranking studies have been used by P&T and hiring committees to try and gauge the worth of a researcher. Some researchers may be too junior to even show up on the researcher rankings. This fact does not exclude them from the practice of rankings. Journal rankings are still used to see the worth of the publications of a junior researcher. The publication list of a researcher, no matter how short, is scrutinized by venue that the research has been published in. This even goes outside of journals as conference proceedings are also ranked according to prestige, difficulty of acceptance, and worthiness in the field.

Most ranking studies focus in on the journal, some focus in on the institution and yet others focus in on the individual researcher (Alexander et al. 2007; Athey and Plotnicki, 2000; Baskerville, 2008; Chua, Cao, Cousins, and Straub, 2002; Clark and Wright, 2007; Ferrat, Gorman, Kanet, and Salisbury, 2007; Geary, Marriott, and Rowlinson; 2004, Hardgrave and Walstrom, 1997; Harzing, 2008; Huang and Hsu, 2005; Korobkin, 1999; Kozar, Larsen, and Straub, 2006; Liang, 2006; Lowry et al., 2004; Martin, 2007; Mingers and Harzing, 2007; Mylonopoulos and Theoharakis, 2001; Nelson, 2006; Nerur and Sikora, 2005; Peffers and Tang, 2003; Podsakoff and MacKenzie, 2005; Rainer et al., 2005; Willcocks, Whiley, and Avgerou, 2008). For the purpose of this paper, 'ranking studies' will refer to as the whole of these three types of studies.

JOURNAL/RESEARCHER RANKING STUDIES IN IS

Ranking studies in the past have typically included two types. One was the survey that was sent out to various researchers in the field to assess what they thought of the worthy journals, institutions, or researcher (Hardgrave and Walstrom, 1997). This type of study tries to get a grasp of the pulse of the field. By asking the very people in the field the survey study is giving voice to those that make up the population. The motivation is honest and unless you ask the populous you cannot find out what the people think. The populous include committee members and are thus the ones that are making hiring and firing decisions based on what they think are good journal rankings so what the populous thinks is valid to the field.

The problem in survey studies are that they are engrossed in subjectivity and what the populous thinks may not be always what the statistics indicate. Survey studies are limited in nature by the fact that surveys cannot incorporate every avenue that research can be published. Typically survey studies list out journals for the surveyee to choose from. The survey may also indicate to the surveyee that they need to 'write in' any journal outlets that may have been omitted from the survey. This has been shown to cause an 'anchor' effect where journals that are listed are more likely to be ranked higher than those that are not listed. Surveyees may also have their own motives to rank a journal/institution/researcher in a particular way. Perhaps they are a high-ranking editor at a journal. Then they would be motivated to push their journal higher up in the rankings. Perhaps they have had multiple publications at a particular journal. Past publications would motivate them to rank that particular journal higher up. Perhaps they were rejected or had a bad experience with a journal. This would motivate them to rank the particular journal lower. One can assume that in institutional rankings, people are motivated to rank their own home institution higher than other institutions that they do not have ties to. Where they work, where they worked, or their alma mater will probably garnish favorable rankings (Alexander et al. 2007; Baskerville, 2008; Peffers and Tang, 2003; Powell and Woerndl, 2008; Rainer et al., 2005).

The second type of ranking studies tries to take a less subjective approach and try to come from the statistical point of view (Lowery et al. 2007). These studies typically have looked at the citation counts and use this as a surrogate for academic worth. By looking at citations the study is looking purely at statistics. But subjectivity is still incorporated in the citation. A citation is when someone cites a previously published paper. Thus some researcher thought that what was said in the cited paper was worth academic credence, and thus the paper is cited. Here lies the subjectivity in this type of study.

The problem with citation analysis is that the culture of the way a field cites can affect the results (Galliers and Whitley, 2007; Gallivan and Benbunan-Fich, 2007; Whitley and Galliers, 2007). If a journal has a limit to the number of citations, citation counts are artificially being suppressed, thus worthy works that could get citations are not getting the recognition they deserve. In the opposite direction citation padding and self-citation practices can cause unworthy papers to get more citations. A researcher may be inclined to boost their own citations counts so they may add their past research to the current research just for the sake of increasing their citation count. Reviewers may encourage a submitter to include the reviewers' previous works thus padding the reviewer's citations. Journal reviewers may indicate that their own papers be cited in the review process. Citations also don't indicate a strength or direction. For example if a current paper draws multiple ideas from a previous paper and multiple citations are made to the same previous paper throughout the current paper, the previous paper is only credited with one citation. Citations may also be saying something negative about the previous paper. For example if a paper is refuting a previous work and saying that the previous work is flawed and wrong, the citation is made but the negative manner is not recorded. That is still one citation, whether positive or negative, for the previous paper.

Both types of studies are wrought with time-consuming research. Surveys can take months or maybe years to complete. Citation analysis studies include the use of knowledge basis like Thompsons ISI in order to get data and then careful analysis. In either case, there is a call for ranking studies to be repeated (Rainer et al. 2005; Lowery et al. 2007). The reason for this is that the knowledge creation field is a living, constantly changing area. Some journals may fall out of top rankings. Falling may be due to the degradation in quality of the papers presented, but not necessarily so. A focus change may cause a journal to fall from the ranks of an academically motivated study, but may increase in rank for a practitioner motivated study.

BIBLIOMETRIC TOOLS

Recent advances in computing technology and the ability to calculate quickly have allowed new and more complex bibliometric tools to be developed and used. Past citation analysis may have been done by hand, or even if computers were used, data collection and analysis may have been done on a spreadsheet or special computer program modules may have been created. There are a few new bibliometric tools that have started to be noticed by the library sciences community and the reach has been to IS as well (Egghe, 2006; Hirsch, 2005; Sidiropoulos, Katsaros, and Manolopoulos, 2006). One of these bibliometric tools are the Hirsch family of indices.

H-INDEX

The H-index is a measure of productivity of a researcher or group of researchers. The h-index tries to give a measure of the worth of a researcher by taking a measure of productivity and quality. Productivity is measured by using the number of papers that the author has produced (Hirsch, 2005). The citation is used as a congregate to quality. The author realizes that quality cannot be solely measured by citations. This is a limitation to the h-index. The problem lies in the spirit of the h-index. The h-index was created in order to have an index that quickly measures the worth of a researcher. The index incorporated quality and productivity. The author agrees that quality can be measured via the actual reading of the work and assignment of some measure of quality. But in order to have an index that is quickly identified, the h-index uses citations.

Citations are the only qualitative measure of a paper to identify influence. By influencing other researchers, citations are giving credit to other researchers. Citations may have positive or negative influence.

Citations may bring positive influence by allowing future researcher to cite a paper due to the positive philosophical expanse that the previous paper has provided. The past paper may have made a mistake or is wrong. Future papers may look at the paper negatively and cite accordingly. The h-index does not differentiate. Either way, positive or negative, the past paper has influence on the future researcher. The discussion of influence and quality is not the focus of this paper.

The h-index is defined as the number of papers h if h of a set of N papers have at least h citations each, and where all other (Np -h) papers have no more than h citations each (Hirsch 2005). In order to get the Hirsch index one needs to get all the papers that the researcher has produced. Then the papers are ranked according to the number of citations that the papers have received in descending order. For example taking a fictional journal, Journal of Indicies (JOI), one need to first rank the papers according to citations (table 1). When the rank overtakes the citation number is where the h-index lies. The table can continue with the complete publication list of JOI but for the purpose of calculating the h-index one only need to get to where the rank overtakes the citation number. One can assume that the remaining papers after rank 9 have 6 or less citations each.

Rank	Atricle Title	Citations
1	The R index	45
2	The Y index	42
3	The A index	35
4	The E index	23
5	The Bc index	15
6	The Or Index	7
7	**The W index**	**7**
8	The Pb index	7
9	The Ri index	6

Table 1. H-index Example

We see that at rank 7 is where the h-index is. We see at rank 6, we have 6 papers that have at least 6 citations each, but the remaining papers have more than 6 citations (the papers at rank 7 and 8 both have more than 6 citations), thus we need to go more. We see that at rank 7, we have 7 papers that have at least 7 citations each, and the remaining papers have no more than 7 papers each. So the h-index condition is satisfied here. Going one more further at rank 8 we do not see 8 papers with at least 8 citations (rank 6, 7, and 8 have only 7 citations each), so the h-index is seven. Two of the problems (other than the quality debate brought up earlier) with the h-index is that extremely large citations have little influence, and the advantageous nature of older aged researchers contribute the weakness of the h-index (Egghe, 2005; Sidiropoulos et al. 2006).

G-INDEX

The G-Index tried to alleviate the problem associated with the extremely large citations that the h-index does not address (Egghe, 2005). When a researcher has a particularly large citation count for a low number of articles, with the h-index, those large citations count article only contributes as much as smaller citation count articles. For example two researchers shown in table 2 have the same h-index (3). But researcher R1 has two articles that have a large amount of citations. One can see that R1 has more influence over the field yet the h-index does not account for the size of the citations.

Rank	R1 Article Title	R1 Citation Count	R2 Article Title	R2 Citation Count
1	The A index	5030	The D index	6
2	The B index	2303	The E index	5
3	The C index	4	The F index	4

Table 2. Two Researchers with H-index of 3

The G-index takes the citation counts into effect by keeping a running total of the citations. The g-index if defined as '*A set of papers has a g-index, g, if g is the highest rank such that the top g papers have, together, at least* g^2 *citations (Egghe 2005,p.*

132).' The G-index is found similar to the h-index in that the rankings are put in order and the number at which the rank squared takes over the running total is your g-index. Taking the same examples as seen when calculating the h-index, we see how the g-index is calculated in table 3. The g-index is found at the point of g equal to 14.

Rank	Rank^2 (g^2)	Atricle Title	Citations	Running Total
1	1	The R index	45	45
2	4	The Y index	42	87
3	9	The A index	35	122
4	16	The E index	23	145
5	25	The Bc index	15	160
6	36	The Or Index	7	167
7	49	The W index	7	174
8	64	The Pb index	7	181
9	81	The Ri index	6	187
10	100	The Yi index	5	192
11	121	The Ar index	4	196
12	144	The Au index	3	199
13	169	The Bo index	3	202
14	**196**	**The Cr index**	**2**	**204**
15	225	The Sn index	1	205
16	256	The Al index	1	206

Table 3. G-index Example

One of the main problems with the g-index is that like the h-index, the age of the researcher or journal is still a problem (Sidiropoulos et al. 2006).

HC-INDEX

The hc-index tries to alleviate the problem of age that exists with the h and g indices (Sidiropoulos et al. 2006). Both the h and g-index favor a researcher or journal that has been around for some time over one that is new in existence. Both are tied to citations so a researchers article needs to be in existence for some time in order for the article to be cited. Also, the longer a researcher article is in existence the more likely that the article will get citations. So naturally a new article has little chance of gaining a large number of citations compared to an article that has been in existence for a longer time period. The hc-index uses a formula to compensate for younger articles that are getting citations and older articles that have established citations. The formula is:

$$S^C(i) = \frac{\gamma * C(i)}{(Age(i)+1)^{\delta}}$$

The i is the current time. C(i) is the citation that a paper gets at time i. The Age(i) is the age of the paper in years. A current year paper would get an Age(i) of zero. The γ is a compensation factor that allows younger papers to get more recognition as the γ is increased. The δ is a compensation factor that allows older papers to get less recognition as the δ is increased. Originally Sidiropoulos et al. (2006) used $\gamma = 4$ and $\delta = 1$, which is what is used in the current research. The hc-index is then calculated similar to the h and g indices. The $S^c(i)$ score is used to rank the papers, similar to the h-index. Then when the rank overtakes the $S^c(i)$ score is where one finds the hc-index. Using the same running example we find:

Rank	Atricle Title	Age	Citations	$S^c(i)$
1	The E index	1	23	46.00
2	The A index	8	35	15.56
3	The Ri index	1	6	12.00
4	The R index	16	45	10.59

5	The Y index	15	42	10.50
6	The Or Index	2	7	9.33
7	**The Bc index**	**6**	**15**	**8.57**
8	The Ar index	2	4	5.33
9	The Pb index	8	7	3.11

Table 4. hc-index Example

Table 4 shows the rankings shows the article title, the age of the article in years, the number of citations, and the $S^c(i)$ score calculated with the equation above. The list is then ranked according to the $S^c(i)$ score. The hc-index is then found in the similar fashion as the h and g indices where the hc-index is where the rank overtakes the $S^c(i)$ score. Here we see that the hc-index is 7 for our running example. Note that now the highest ranked paper is the paper titled "The E index" which was previously ranked fourth but due to the amount of citations (23) and the relative young age of the paper (1 year old) the paper overtakes the three papers before it to become highest ranked in this list. The previously highest ranked paper "The R index" is now ranked fourth because of its age (16 years). In this fashion the hc-index gives more credence to papers that have citations and are younger in age. The problem with the hc index is that larger hits are not accounted for just like the h-index didn't account for large hits.

INTRODUCTION OF THE GC INDEX

This research attempts to alleviate the problem that the h, g, and hc index by combining the factors that the g and hc-index incorporates and proposes a new index called the gc-index. With this index the $S^c(i)$ is used (as with the hc-index) but a running tab will be kept and compared with the rank squared (as with the g-index). Continuing our running example we have:

Rank	$Rank^2$	Atricle Title	Year/Age	Citations	$S^c(i)$	$S^c(i)$ Total
1	1	The E index	1	23	46.00	46.00
2	4	The A index	8	35	15.56	61.56
3	9	The Ri index	1	6	12.00	73.56
4	16	The R index	16	45	10.59	84.14
5	25	The Y index	15	42	10.50	94.64
6	36	The Or Index	2	7	9.33	103.98
7	49	The Bc index	6	15	8.57	112.55
8	64	The Ar index	2	4	5.33	117.88
9	81	The Pb index	8	7	3.11	120.99
10	**100**	**The Yi index**	**6**	**5**	**2.86**	**123.85**
11	121	The W index	9	7	2.80	126.65
12	144	The Bo index	7	3	1.50	128.15

Table 5. gc-index Example

By keeping a running total of the $S^c(i)$ one is able to find the gc-index in a similar fashion as the h, g, and hc-indices where the $rank^2$ is overtaken by the $S^c(i)$. We find the gc-index for our running example to be 10.

RESULTS

Three journals from North America and three journals from Europe were chosen according to Rainer and Miller (2005). These were *MIS Quarterly* (MISQ), *Information Systems Research* (ISR), and *Journal of Management Information Systems* (JMIS), the *European Journal of Information Systems* (EJIS), the *Journal of Strategic Information Systems* (JSIS), and the *Information Systems Journal* (ISJ). The data was retrieved during October of 2008 from Google Scholar (GS). GS has been criticized and hailed as a possible means of retrieving citation analysis data. The argument of whether GS is a good data source is outside the scope of this paper. After the data was retrieved the g, h, hc, and gc index were calculated. The results are in table 6, and figure 1. There were no changes in the rankings of the journals according to any of the indices.

	h	g	hc	gc
MISQ	136	233	76	111
ISR	90	160	52	85
JMIS	84	134	47	72
EJIS	47	74	28	41
JSIS	41	63	26	37
ISJ	42	65	27	37

Table 6. Results by Journal

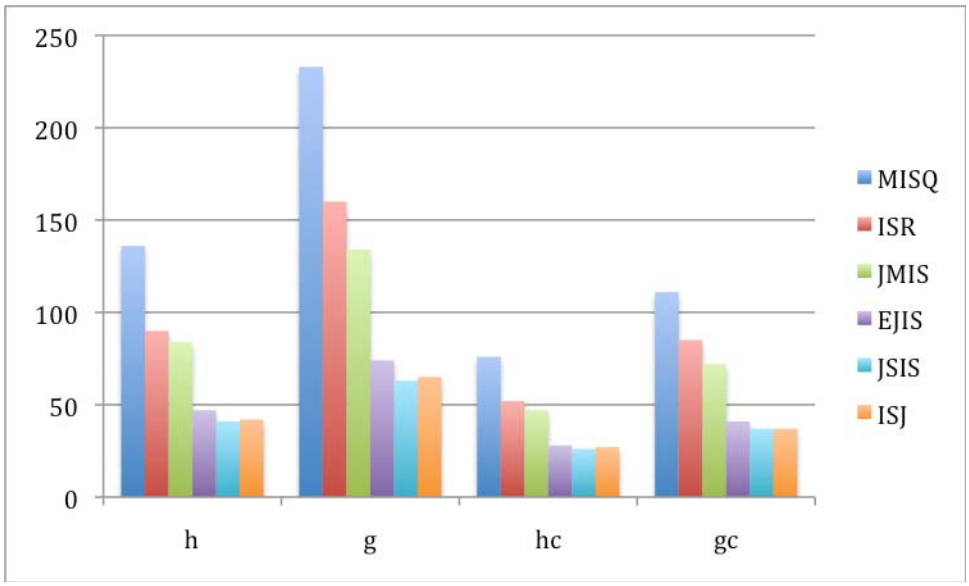

Figure 1. Comparison of indices

When looking at the chart we see that the g-index shows a big disparity between MISQ and the other journals while the hc-index shows less and almost no difference between ISJ and JMIS. With the gc-index we see that there is a difference but the lead by MISQ is understated compared to the g-index and the difference between ISR and JMIS is accentuated compared to the hc-index.

These differences are not necessarily a feature of the gc-index and can be a function of the dataset. The gc-index combines the features of the g and hc-indices and thus the results showing a spread may be the function of the gc-index.

CONTRIBUTIONS AND LIMITATIONS

The use of the gc-index is not an end all for journal/researcher ranking. The gc-index is proposed as a measure to add to the assessment of a body of research. The use of the Hirsch family of indices, as well as other assessment methods (quality of papers, reputation), should be used in addition to the gc-index.

The limitations of this research, as well as future research areas are two fold. First, this research is a proposal for the use of the gc-index. The gc-index needs to be tested on more datasets. Second, the dataset was taken from GS, which may not be the most reliable database for taking data.

CONCLUSION

This paper proposes a new metric for assessing of a body of research. The gc-index was introduced and used for assessing six major IS journals. The gc-index was seen as combination of the g and hc-indices. Although the ranking of the six journals didn't change for the gc-index, the dataset showed that gc-index was different from the other indices.

REFERENCES

1. Alexander, J.K., Scherer, R.F., and Lecoutre, M. "A Global Comparison of Business Journal Ranking Systems," Journal of Education for Business (82:6), July/August 2007, pp 321-327.
2. Athey, S., and Plotnicki, J. "An Evaluation of Research Productivity in Academic IT," Communications of the Association for Information Systems (3:7) 2000, pp 2-19.
3. Baskerville, R. "Introduction: For better or worse: how we apply journal ranking lists," European Jounral of Information Systems (17) 2008, pp 156-157.
4. Chua, C., Cao, L., Cousins, K., and Straub, D. "Measuring Researcher-Production in Information Systems," Journal of the Association of Information Systems (3) 2002, pp 145-215.
5. Clark, T., and Wright, M. "Reviewing Journal Rankings and Revisiting Peer Reviews: Editorial Perspectives," Journal of Management Studies (44:4), June 2007, pp 612-621.
6. Egghe, L. "Theory and practice of the g-index," Scientometrics (69:1) 2006, pp 131-152.
7. Ferratt, T.W., Gorman, M.F., Kanet, J.J., and Salisbury, W.D. "IS Journal Quality Assessement Using the Author Affiliation Index," Communications of the Association for Information Systems (19) 2007, pp 710-724.
8. Galliers, R.D., and Whitley, E.A. "Vive les differences? Developing a profile of European information systems research as a basis for international comparisons," European Journal of Information Systems) 2007.
9. Gallivan, M.J., and Benbunan-Fich, R. "Analyzing IS research productivity: an inclusive approach to global IS scholarship," European Journal of Information Systems (16:1) 2007.
10. Geary, J., Marriott, L., and Rowlinson, M. "Journal Rankings in Business and Management and the 2001 Research Assessment Exercise in UK," British Journal of Management (15) 2004, pp 95-141.
11. Hardgrave, B.C., and Walstrom, K.A. "Forums for MIS Scholars," Communications of the ACM (40:11) 1997, pp 119-124.
12. Harzing, A.-W. "On becoming a high impact journal in international business and management," in: European Journal of International Management, 2008.
13. Hirsch, J.E. "An index to quantify an individual's scientific research output," Proceedings of the National Academy of Sciences of the United States of America (102:46) 2005, pp 16569-16572.
14. Huang, H.-H., and Hsu, J.S.-C. "An Evaluation of Publication Productivity in Information Systems: 1999-2003," Communications of the Association for Information Systems (15) 2005, pp 555-564.
15. Korobkin, R. "Ranking Journals: Some Thoughts on Theory and Methodology," Florida State University Law Review (26) 1999, pp 851-876.
16. Kozar, K.A., Larsen, K.R., and Straub, D. "Leveling the Playing Field: A Comparative Analysis of Business School Journal Productivity," Communications of the Association for Information Systems (17) 2006, pp 524-538.
17. Liang, L. "h-index sequence and h-index matrix: Constructions and applications," Scientometrics (69:1) 2006, pp 153,Äì159.
18. Lowry, P.B., Karuga, G.G., and Richardson, V.J. "Assessing Leading Institutions, Faculty, And Articles in Premier Information Systems Research Journals," Communications of the Association for Information Systems (20) 2007, pp 142-203.
19. Martin, N. "Keeping Score: CJR Presents a New Paradigm in Rating Journals," Econtent (30:3), April 2007, p 14.
20. Mingers, J., and Harzing, A.-W. "Ranking Journals in Business and Management: A Statistical Analysis of the Harzing Dataset," in: European Journal of Information Systems, 2007.
21. Mylonopoulos, N.A., and Theoharakis, V. "Global Perceptions of IS Journals," Communications of the ACM (44:9) 2001, pp 29-33.
22. Nelson, M.J. "An Alternate Method of Ranking Journals using Citations," University of Western Ontario, London, Ontario, Canada, 2006, p. 7.
23. Nerur, S., and Sikora, R. "Assessing the Relative Influence of Journals in a Citation Network," Communications of the ACM (48:11) 2005, pp 71-74.
24. Peffers, K., and Tang, Y. "Identifying and Evaluating the Universe of Outlets for Information Systems Research: Ranking the Journals," The Journal of Information Technology Theory and Application (5:1) 2003, pp 63-84.

25. Podsakoff, P.M., MacKenzie, S.B., Bachrach, D.G., and Podsakoff, N.P. "The influence of management journals in the 1980s and 1990s," Strategic Management Journal (26:5) 2005, pp 473-488.
26. Powell, P., and Woerndl, M. "Opinion Paper: Time to stop researching the important things?," European Jounral of Information Systems (17) 2008, pp 174-178.
27. Rainer, J., R. Kelly, and Miller, M.D. "Examining Differences Across Journal Rankings," Communications of the ACM (48:2) 2005, pp 91-94.
28. Sidiropoulos, A., Katsaros, D., and Manolopoulos, Y. "Generalized h-index for Disclosing Latent Facts in Citation Networks," arXiv:cs.DL/o606066 (1), 07/13/2006 2006.
29. Whitley, E.A., and Galliers, R.D. "An alternative perspective on citation classics: Evidence from the first 10 years of the European Conference on Information Systems," Information & Management (10:2) 2007.
30. Wilcocks, L., Whiley, E.A., and Avgerou, C. "Opinion Paper: The ranking of top IS journals: a perspective from the London School of Economics," European Jounral of Information Systems (17) 2008, pp 163-168.