

7-1-2009

Journal Self-Citation XXI: Bibliography as Artifact – How Citations Are Data

Eleanor Wynn

Intel Corporation, Eleanor.Wynn@intel.com

Follow this and additional works at: <https://aisel.aisnet.org/cais>

Recommended Citation

Wynn, Eleanor (2009) "Journal Self-Citation XXI: Bibliography as Artifact – How Citations Are Data," *Communications of the Association for Information Systems*: Vol. 25 , Article 21.

DOI: 10.17705/1CAIS.02521

Available at: <https://aisel.aisnet.org/cais/vol25/iss1/21>

This material is brought to you by the AIS Journals at AIS Electronic Library (AISeL). It has been accepted for inclusion in Communications of the Association for Information Systems by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

Communications of the Association for Information Systems

CAIS 

Journal Self-Citation XXI: Bibliography as Artifact – How Citations Are Data

Eleanor Wynn

Information Technology Research, Intel Corporation

Eleanor.Wynn@intel.com

Abstract:

Citation rates are important criteria for judging an author's impact in the research community or discipline. They form one of the many indicators of research quality that help people less familiar with a field understand the effectiveness of a research paper or individual. The contents of a bibliography or set of references are also markers that help orient the reader to the framework of a piece of writing. Lately, journal publishers have paid increased attention to citation rates as more and more journals vie to be included in abstracting and indexing services. These lists act as a filter for libraries and bulk subscription brokers, so they can affect circulation and hence, revenue. Because of this, some journals have required authors to cite prior papers from the same journal when submitting research, even going so far as to specify how many citations are needed for a paper to be considered. This paper argues that commercial considerations must not outweigh the primary academic purpose of a citation list. Not only are there a host of conventions associated with citations that arise from author integrity and community relationships, but the bibliography is itself research data for another discipline. Compelling and precise portraits of researchers can be derived from such data. A case from library science on the late Rob Kling's sources illustrates the value of preserving the integrity of citations for the purpose of building and understanding a discipline, not a revenue stream.

Keywords: citation rates, research impact, library science

Volume 25. Article 21. pp. 171-176. Julv 2009

I. INTRODUCTION

In February 2009, this message was posted to the AISWorld e-mail distribution list:

I received the following e-mail from a [Major Publisher] journal on whose editorial board I serve:

"The EICs are requesting that henceforth all articles that are accepted for publication to <journal x> should cite at least five <journal x> articles. This is common practice for all top journals."

I would appreciate input from my colleagues about this message.

Is this really a common practice?

Is it appropriate to reference the publishing journal over and over whether it is necessary to the paper or not?

Is it ethical?

The query elicited 92 essay responses of a paragraph or more that were either published to the list or sent to the message's author. There were also shorter yes/no responses to each of the questions [Gray 2009]. Some of the responses against the practice were outraged, while some mildly in favor conveyed a spirit of laissez-faire. One contributor proposed to use this citation practice as the basis for a new metric: self referencing vs. diversely referenced journals. This is a Solomonic approach as it uses the self referencing behavior as an indicator in its own right; that would also make it self defeating.

I was one of those opposed to required citations on the grounds that AEs and reviewers ought to be able to determine the paper's fit with a journal without having the journal explicitly referenced. The focus should be on the body of literature cited. That body of literature ought to have some familiar touch points that reflect the journal's mission and core audience. But it does not have to come directly from the journal itself. If there is no familiar literature context in the paper, then the points may miss the audience. It if happens that the journal in question had a pertinent article on the topic, it could be pointed out, but should not be a filtering mechanism for the paper's acceptance. There may be a little shiver of disappointment that one's own journal was overlooked by someone submitting. But authors do not submit to a journal based on being its devoted followers. They, too, look for a good fit. When they haven't done so, it is usually blatantly obvious.

I also felt, in response to the last of the questions, that the practice of requiring references to a specific journal is not ethical because it tampers with the journal's citation index accuracy, a topic I am acquainted with as a former editor in chief. The publishers want to get their journals into abstracting and indexing services that convey status and bring in revenues. These services require a certain level of citations as a condition of including the journal in their lists. That is a worthy goal. It sells journals to more libraries. There is a need from the academic side to publish in better ranked journals; this need overlaps with the publishers' need to have their journals sell and attract known authors. Authors like to be widely read and, of course, cited. The mutual interests intersect at many points. But when citations directly become a marketing metric, a critical line is crossed. That is the line between the commercial interest of the publisher and the academic standards of the disciplines, be they sciences, humanities, a mix, or undecided.

The paper reviews the well known purpose of the bibliography in supporting research papers. The bibliography lends credibility and saves many pages of digression. However there is another key purpose that goes deeper: bibliographies are themselves research data. They are used to trace the provenance of concepts and to trace the influence of key researchers. They paint a portrait both of a discipline and of the people in it.

II. THE PURPOSE OF A BIBLIOGRAPHY

Why should the citations in an article be a matter of academic integrity? Let us review their purpose. Others who replied to the e-mail thread expressed this value eloquently. It is a much visited subject in every discipline, not to mention every research writing class.

Literacy

Most people who submit journal articles attended graduate school. Most graduate schools require students to read extensively in an area of specialization. Students are examined on the literature in this specialization. Presumably, the dissertation and subsequent articles and books will also be based in this specialization or use it as a starting point. Readers want to know who your sources are, whether you have shotgun references, and how well you understand the literature you refer to. It has become a common practice to skim through any references that might suit the point and use those as indicators without showing a deep knowledge of the sources. However, in most of the source disciplines for Information Systems, especially the social sciences and most particularly sociology, the literature review is supposed to be an in depth discussion of the prior arguments on the topic, knowledge of which led the author to conclude that there is a gap or point of disagreement. This gap or disagreement, or additional data, are the purpose of publishing an article.

Rigor

Rigor is a term that has been abused in IS, along with sociology. It is often applied to the mathematical and statistical values of a piece of research unevenly with the theoretical, plausibility, or relevancy values. I don't mean that kind of rigor. What I mean by rigor in a literature review is the demonstration that an author seriously considered all sides of a position, then systematically dispensed with the ones disagreed with and parsed the ones agreed with. This rigor is what makes room for the research at hand. Because some past work is demonstrably deficient or misguided and because other research covered some different points or missed a new development, it is therefore important to write the present article to set the record straight. That is also the reason why a journal should publish this important article. Literature review rigor makes for topic relevance and situatedness in an ongoing discourse as neither redundant nor uninformed.

Fairness

The title of Robert K. Merton's classic work, *On the Shoulders of Giants* [1993], has become a shorthand phrase for the endeavor of trying to develop theory in any domain. Even Kepler relied on earlier sources [Kuhn]. In this century, we are most certainly in a position of adding a small amount to a large body of existing discussion. In IS we are fortunate that the domain changes so rapidly that it is easy to make a novel contribution just by researching a topic of the day. But new work ought to have some basis in prior theory and prior exploration of someone else's data, or at the very least recognition of who else has worked on this topic in the past few years. It is ethical to recognize those others. Not doing so risks a charge of plagiarism. Stealing the work of others is possibly the worst crime in academia.

Courtesy and Membership

Aside from the question of fairness as an ethical principle, there is the simple matter of professional courtesy and recognition that a community exists. All participants depend on this community for their own recognition, for resources, employment, and references. People run into each other at conferences and recognize each other from citations, or from not having been cited. A bibliography represents a social network of common intellectual bonds.

A literature can also divide people into groups across theoretical and methodological boundaries. This is not a bad thing; it is efficient. Some people start to read a paper at the back, at the references, as a filtering step. This approach is a form of social recognition that requires only a few cues to create a sense of sameness or otherness. Most people want these cues about themselves to be both presentable and accurate. Although requiring a certain reference may not ruin an image, it might. What if the only articles found in the journal in question were ones to disagree with? Would that matter? If count is all that matters, what is actually said about the citations might be moot until the policy becomes even more restrictive.

In short, much information is given by bibliographies. They form a function within the context of a discipline that transcends the needs of a publisher to increase the citation index value for a journal.

III. THE USES OF A BIBLIOGRAPHY

The last section gave a short overview of the intellectual and social contexts of a bibliography. I am not an expert on this subject, but my eyes were opened by two talks I heard at the memorial conference for Rob Kling held at UC Irvine, March 11-12, 2005. The two speakers [Cronin and Shaw, Robbin] were distinguished scholars of library science and colleagues of Rob at Indiana University. I had not seen the high art of library studies until that day. Both authors read their papers from the manuscripts as is still done in some disciplines, mainly the traditional ones where neither math formulas, nor demos, nor PowerPoint slides are remotely adequate to the purpose of developing the fine points of an idea.

These two papers (and the citations they referred me to) are convincing evidence that a bibliography is an artifact that ought not, under any circumstances, be tampered with for purposes other than documenting an author's intellectual and factual sources. The bibliography traces what formed the author. Once the person has a body of work, bibliographic research traces what and who that author has influenced. To insert content coercively is to cover the path backwards or to place false clues on that path. I am not speaking of the case where someone sees an omission and gladly corrects it, but the case where five citations from a particular journal are required in order to be considered for publication. If this were a universal practice, an author would appear to be a grazer wandering from journal to journal for key sources rather than a pursuing a consistent thread, or at least being a self determined widely read grazer.

Cronin and Shaw [2007] refer to the term "citation identity" in their paper reviewing the sources of Rob's intellectual formation:

"Rob Kling was a singular scholar, but he was no singularity." Thus begins an excellent review of where Rob's seminal ideas and formation came from and how he developed them into his own perspective...As Shapin [1995; p. 359] demonstrated in his analysis of the political and moral economy of the seventeenth century science in England, "scientific work is produced by and in a *network* of actors." That is no less true today. [Cronin and Shaw]

Another key term in that paper, "cognitive microenvironments," comes from Merton [1988]. Just as evocatively, the authors cite Latour's [to be cited] "semiotic actors presented in the text but not *present* in the flesh." Thus, a writer's own citations become a signature for the entire formative context in which new works are produced.

It is true that any journal articles that Rob might have cited for reasons of getting accepted (an unlikely scenario) would have figured in the noise of his overall formation. Nevertheless, were this practice to become systemic, it would introduce some odd distortions in the kind of record the library scholars so painstakingly built of one man's influences. They present several tables of enviable citation rankings in Rob's work for periods of time, of which Rob himself always ranked at the top. Congratulations are in order to Kristen Eschenfelder of the University of Wisconsin-Madison, who was quite junior at the time of Rob's death, who nonetheless made the list for 1974-2005! And who knew Rudy Hirschheim (University of Houston) was so well appreciated by Rob?

Robbins's paper for the memorial was less focused on citation itself and more focused on theoretical formation; less contemporary influences of the network and more historical influence. It is drawn from the same bibliography but has a different focus. So the bibliography itself is a data set that two different people, colleagues in the same discipline, have used to draw two different, but factually based portraits of the same individual.

Not only is there a practice of analyzing bibliographies for intellectual history and influence of an individual but there are of course, *methods*, summarized in the term *bibliometrics*. Following is an example of a terminology, ego-centric citation analysis, its constructs, and of course, its own citations:

The notion of ego-centric citation analysis was introduced by Howard White [2000; 2001a]. White defined the following constructs:

- Coauthors: authors (with multiplicity) who published with the specific author
- Citation identity: authors (with multiplicity) who are cited by the specific author
- Citation image makers: authors (with multiplicity) who cite works of the specific author
- Citation image [White and McCain 1998]: authors (with multiplicity) who have been co-cited with the specific author [Bar-Ilan 2006]

Any field has its rigor and standards of credibility. In some fields, like philosophy, arguments are more important. But in most fields, some form of data are important, whether they be historical records, observational data of naturally occurring events, or experimental/survey data. Bibliographies are data for bibliometrics. And this field can inform us in very specific ways beyond the particular lens we use in our own rather diverse discipline of information systems. Information science can take any discipline and dissect it for its history, sources, and intellectual clusters. All bibliographies are data for this field. Other kinds of analysis include tracing of co-author networks, cross disciplinary analysis to trace innovation trends and intellectual network hubs [cites to come Rhoten].

Even under the current circumstances where we can presume that deliberate citation inflation is not yet widespread, there are problems of data cleansing for bibliometric studies [Bar Ilan 2006, p. 1556]. Bar-Ilan raises two instances:

(1) relating to which of more than one paper an author may have published in a single journal issue, as first or second author, and (2) the other relating to whether or not the indexing service includes second authors in their index and at what point in time they may or may not have done so. Different indexing services have different criteria for what to index, e.g., a paper in a top journal, a paper in a mid or above journal, a lecture or lecture notes, or proceedings papers.

Again using the example of Rob Kling, Bar-Ilan cites from Cronin and Shaw [2002] that certain figures span disciplines or migrate from one discipline to another. This too, becomes part not only of the individual's trace, but of the history of interdisciplinary work and critical interventions.

One of the professors studied, Robert Kling, "migrated" to information science from computer science and management. This is clearly reflected in his citation image makers list, where 73 percent of the citations his works received were outside information science. They also compared the top citation identities with the top image makers lists, and found a relatively small number of names in the intersection of these sets.

IV. CONCLUSION

The foregoing section dabbles in a discipline I don't know well. My limited exposure to it has been revelatory. It adds legitimacy to the general sense that forcing citations is not right, not "done" for ethical reasons and general scholarly culture. But these few readings also support a thesis that a bibliography is as sacred as any other form of data, maybe more sacred because the analysis of it is by nature longitudinal, whereas a survey's findings can be challenged by a new survey.

Even electronic records owned by one individual can be fluid and inconsistent. Consider two recent cases internal to Intel Corporation using experimental methods. In one, we found that in order to validate an agent based model that would indicate when new servers should be acquired, we had to perform massive data cleansing on the prior year records of distributed server usage [Bonabeau, Tan, and Dusevic and Wynn, in preparation]. Naming conventions tended to drift across the different server farms and to diverge historically, along the lines argued by Ciborra [2000]. In another study, machine learning programs tracked task-switching and information re-use of day to day PC usage for as long as two months. [Stumpf et al; Jensen, Wynn and ?? in progress].

In the latter study, tracking provenance of information from one document to another, we found that reuse is pervasive. While it is possible to track that information from one source to another, it is not normally done. Information workers tend to assemble available information into new summaries and pull from where they can find sources. This led us to wonder about the robustness of the average workday "document" as being a thing at all, except for its use as either a repository or a form of rhetoric. Without tracking, users end up with temporary information compilations that move from one application to another in bits and pieces. As a result, it is almost impossible to know the age and consistency of information in documents, the quality of sources, their compatibility with each other, and the kind of attribution that the academic community so relies on.

In short, keeping track of provenance is a non-trivial problem regardless of where we look. Citation conventions arose precisely because of this kind of circumstance in earlier eras and other media.

While the introduction of small distortions in the stream of any data set may seem harmless, in fact we don't know what the effect of that distortion will be to the overall record, especially if the trend amplifies. It is important as well for the individual author to know whether his/her citation index reflects genuine interest in the work or simply reflects the practices of a publisher. Puffing the citations is the equivalent of doping in sports. Suddenly the people who play fair, who rely on natural skill and discipline, are at a disadvantage.

Ethics and appropriateness are not the only considerations against coerced citations. Citations are artifacts that anchor the key theories and facts of a paper and, collectively, the lifetime intellectual orientations of the author. We don't allow product placements in our reputable scholarly papers and we should not allow them in our bibliographies.

REFERENCES

Editor's Note: The following reference list contains hyperlinks to World Wide Web pages. Readers who have the ability to access the Web directly from their word processor or are reading the paper on the Web can gain direct access to these linked references. Readers are warned, however, that:

1. These links existed as of the date of publication but are not guaranteed to be working thereafter.
2. The contents of Web pages may change over time. Where version information is provided in the References, different versions may not contain the information or the conclusions referenced.
3. The author(s) of the Web pages, not AIS, is (are) responsible for the accuracy of their content.

4. The author(s) of this article, not AIS, is (are) responsible for the accuracy of the URL and version information.

Akrich, M. and B. Latour (1992), "A Summary of a Convenient Vocabulary for the Semiotics of Human and Nonhuman Assemblies," in Wiebe E. Bijker and John Law, eds., *Shaping Technology/Building Society: Studies in Sociotechnical Change*, Cambridge, MA.: MIT Press, pp. 259-264.

Bar-Ilan, J. (2006) "An ego-centric Citation Analysis of the Works of Michael O. Rabin Based on Multiple Citation Indexes," *Information Processing & Management*, 42(6), pp.1553-1566.

Ciborra, C. (2000). *From Control to Drift: the Dynamics of Corporate Information Infrastructures* Oxford University Press.

Cronin, B. (2002). "Hyperauthorship: A Postmodern Perversion or Evidence of a Structural Shift in Scholarly Communication Practices?" *Journal of the American Society for Information Science and Technology* 52(7), pp. 558–569.

Cronin, B., and D. Shaw (2002). "Identity-Creators and Image-Makers: Using Citation Analysis and Thick Description to Put Authors in Their Place," *Scientometrics*, 54(1), pp. 31-49.

Cronin, B. and D. Shaw (2007). "Peers and Spheres of Influence: Situating Rob Kling," *The Information Society* 23(4), pp.221-233.

Dragunov, A. N., T. G. Dietterich, K. Johnsrude, M. McLaughlin, L. Li, and J. L. Herlocker (2005). "TaskTracer: A Desktop Environment to Support Multi-tasking Knowledge Workers," *International Conference on Intelligent User Interfaces (IUI-2005)* ACM Press, pp. 75-82.

Gray, P. (2009). "Journal Self-Citation I: Overview of the Journal Self-Citation Papers – The Wisdom of the IS Crowd," *Communications of the Association for Information Systems* 24(1), Article 1, pp. 1-10.

Jensen, C., Lonsdale, H., Wynn, E. and N. Dedual (2009). "Cutting and Pasting Up: 'Documents' and Provenance in a Complex Work Environment," pending acceptance at the 2009 Hawaii International Conference on System Sciences.

Kuhn, T. *The Structure of Scientific Revolutions* University of Chicago Press.

Merton, R. K. (1965; 1985; 1993). *On the Shoulders of Giants* University of Chicago Press.

Rhoten, D. (2007). "The Dawn of Networked Science," *Chronicle of Higher Education*, <<http://chronicle.com/weekly/v54/i02/02b01201.htm>>

Stumpf, S., E. Fitzhenry, and T. Dietterich (2007). "The Use of Provenance in Information Retrieval," *Workshop on Principles of Provenance (PROPR)*.

ABOUT THE AUTHOR

Eleanor Wynn is a principal engineer in Information Technology Research at Intel Corporation. She earned a Ph.D. in Linguistic Anthropology from the University of California at Berkeley. Recruited to Xerox Palo Alto Research Center as a doctoral intern, she moved from the study of communication in hierarchies to the study of situated knowledge in everyday workplace communication. The encounter with computer science that PARC provided led her to challenge then-current computing frameworks and to pursue the study of organizational reasoning about information technology and technological rationales about organizations. She is editor emerita of *Information Technology & People*, having served as editor in chief from 1985 until 2008. Her early entry into the information systems community was a fortuitous context for her work's purpose. She has also published and taught within the IS discipline.

Copyright © 2009 by the Association for Information Systems. Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and full citation on the first page. Copyright for components of this work owned by others than the Association for Information Systems must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, or to redistribute to lists requires prior specific permission and/or fee. Request permission to publish from: AIS Administrative Office, P.O. Box 2712 Atlanta, GA, 30301-2712, Attn: Reprints; or via e-mail from ais@aisnet.org.



EDITOR-IN-CHIEF
 Ilze Zigurs
 University of Nebraska at Omaha

AIS SENIOR EDITORIAL BOARD

Guy Fitzgerald Vice President Publications Brunel University	Ilze Zigurs Editor, CAIS University of Nebraska at Omaha	Kalle Lyytinen Editor, JAIS Case Western Reserve University
Edward A. Stohr Editor-at-Large Stevens Institute of Technology	Blake Ives Editor, Electronic Publications University of Houston	Paul Gray Founding Editor, CAIS Claremont Graduate University

CAIS ADVISORY BOARD

Gordon Davis University of Minnesota	Ken Kraemer University of California at Irvine	M. Lynne Markus Bentley College	Richard Mason Southern Methodist University
Jay Nunamaker University of Arizona	Henk Sol University of Groningen	Ralph Sprague University of Hawaii	Hugh J. Watson University of Georgia

CAIS SENIOR EDITORS

Steve Alter University of San Francisco	Jane Fedorowicz Bentley College	Jerry Luftman Stevens Institute of Technology
--------------------------------------------	------------------------------------	--------------------------------------------------

CAIS EDITORIAL BOARD

Michel Avital University of Amsterdam	Dinesh Batra Florida International University	Indranil Bose University of Hong Kong	Ashley Bush Florida State University
Fred Davis University of Arkansas, Fayetteville	Evan Duggan University of the West Indies	Ali Farhoomand University of Hong Kong	Sy Goodman Georgia Institute of Technology
Mary Granger George Washington University	Ake Gronlund University of Umea	Douglas Havelka Miami University	K.D. Joshi Washington State University
Chuck Kacmar University of Alabama	Michel Kalika University of Paris Dauphine	Julie Kendall Rutgers University	Claudia Loebbecke University of Cologne
Paul Benjamin Lowry Brigham Young University	Sal March Vanderbilt University	Don McCubbrey University of Denver	Fred Niederman St. Louis University
Shan Ling Pan National University of Singapore	Jackie Rees Purdue University	Jia-Lang Seng National Chengchi University	Paul Tallon Loyola College, Maryland
Thompson Teo National University of Singapore	Craig Tyran Western Washington University	Chelley Vician Michigan Technological University	Rolf Wigand University of Arkansas, Little Rock
Vance Wilson University of Toledo	Peter Wolcott University of Nebraska at Omaha	Yajiong Xue East Carolina University	

DEPARTMENTS

Global Diffusion of the Internet Editors: Peter Wolcott and Sy Goodman	Information Technology and Systems Editors: Sal March and Dinesh Batra
Papers in French Editor: Michel Kalika	Information Systems and Healthcare Editor: Vance Wilson

ADMINISTRATIVE PERSONNEL

James P. Tinsley AIS Executive Director	Vipin Arora CAIS Managing Editor University of Nebraska at Omaha	Copyediting by Carlisle Publishing Services
--------------------------------------------	------------------------------------------------------------------------	---------------------------------------------

